Structural and Biochemical Investigation of the *Yersinia enterocolitica* Type III Secretion Export Apparatus

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Ulrich Andreas Philipp Wiesand
aus Füssen

1. Referent: Honorarprofessor Prof. Dr. Dirk Heinz
2. Referent: Professor Dr. Michael Steinert
eingereicht am: 31.03.2010
mündliche Prüfung (Disputation) am: 16.06.2010

Druckjahr 2010

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch den Mentor der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikation

Wiesand, U., Sorg, I., Amstutz, M., Wagner, S., van den Heuvel, J., Lührs, T., Cornelis, G.R. & Heinz, D.W. (2009). Structure of the Type III Secretion Recognition Protein YscU from *Yersinia enterocolitica*. *J. Mol. Biol.* **385**, 854-866.

Tagungsbeiträge

Wiesand, U., Sorg, I., Amstutz, M., Wagner, S., van den Heuvel, J., Lührs, T., Cornelis, G.R. & Heinz, D.W.: Structure of the Type III Secretion Recognition Protein YscU from *Yersinia enterocolitica*. (Poster), VII European Symposium of The Protein Society, Zürich, Switzerland (2009).

Wiesand, U., Sorg, I., Amstutz, M., Wagner, S., van den Heuvel, J., Lührs, T., Cornelis, G.R. & Heinz, D.W.: Structure of the Type III Secretion Recognition Protein YscU from *Yersinia enterocolitica*. (Vortrag), 17. Jahrestagung der Deutschen Gesellschaft für Kristallographie, Hannover, Germany, (2009)

Wiesand, U., Sorg, I., Amstutz, M., Wagner, S., van den Heuvel, J., Lührs, T., Cornelis, G.R. & Heinz, D.W.: Structure of the Type III Secretion Recognition Protein YscU from *Yersinia enterocolitica*. (Vortrag), 4th Student´s Meeting of the Network of Excellence (NoE) EuroPathoGenomics, Mallorca, Spain (2009).

Wiesand, U., Pagel, H., Sorg, I., Cornelis, G.R. & Heinz, D.W.: Crystal Structure of the Cytosolic Domain of YscU – a Type III Secretion Export Regulator from *Yersinia enterocolitica.* (Vortrag) 10th Heart of Europe Bio-Crystallography Meeting, Bedlewo, Poland (2007).

Content

Abbrevations

A	Absorption
Å	Ångstrøm (0.1 nm)
AF4	Asymmetric flow field flow fractionation
AO	Atomic orbital
AU	a) Absorption unit; b) Asymmetric unit
CC	Correlation coefficient
CCP4	Collaborative Computational Project No. 4
CV	Column volume
Da	Dalton
DLS	Dynamic light scattering
DTT	Dithiotheritol
EHEC	Enterohemorrhagic *E. coli*
EM	Electron microscopy
EPEC	Enteropathogenic *E. coli*
Flh	Flagellar homolog
GPC	Gel permeation chromatography
GST	Glutathione S-transferase
HMQC	Heteronuclear multiple quantum coherence
HRP	Horse radish peroxidase
IEC	Ion exchange chromatography
IPTG	Isopropyl-β-D-thiogalactoside
LCAO	Linear combination of atomic orbitals
MAD	Multiple wave length anomalous dipersion
MO	Molecular orbital
M_r	Molecular mass

MR	Molecular Replacement
MWCO	Molecular weight cut off
NMR	Nuclear magnetic resonance
OD_{λ}	Optical density at wavelength λ
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pI	Isoelectric point
R_h	Hydrodynamic radius
r.m.s.d.	Root mean square diviation
rpm	Rounds per minute
SDS	Sodium dodecyl sulphate
SDS-PAGE	SDS-polyacryalmid gel electrophoresis
SeMet	Selenomethione
SLS	Static light scattering
STEM	Scanning transmission electron microscopy
TCEP	Tris(2-carboxyethyl)phosphine
TxS	Type x secretion
TxSS	Type x secretion system
V_e	Elution volume
V_M	Matthews coefficient
Y.	*Yersinia*
Yop	*Yersinia* outer protein
Ysc	*Yersinia* secretion
$YscU^C$	*Yersinia* secretion protein U (aa 211-354)
$YscV^C$	*Yersinia* secretion protein V (aa 322-704)
$YscV^C{}_{356}$	*Yersinia* secretion protein V (aa 356-704)

Zusammenfassung

Das intestinale Pathogen *Yersinia enterocolitica* verwendet für den Transport von Effektoren über die eukariotischen Zellmembranen eine hoch entwickelte Injektionsmaschine, die auch als *injectisome* oder Typ III Sekretions System (T3SS) bezeichnet wird. Sobald das Pathogen mit einer eukaryotischen Wirtszelle, z.B. einer Makrophage, in Kontakt tritt, wird durch das *injectisome* eine Reihe von Effektoren in diese geschleust. Dort greifen die Effektoren in die zellulären Funktionen der Makrophage ein, sodass diese nicht mehr in der Lage ist das Pathogen zu phagozytieren. Die hochkomplexe molekulare Struktur des T3SS, welche auch in anderen Gram-negativen Pathogenen wie *Salmonella* spp., *Shigella* spp. und enteropathogenen *Escherichia coli* zu finden ist, ist evolutionär mit dem bakteriellen Flagellum verwandt, welches den Proteinexportkanal für den Aufbau seiner äußeren Komponenten nutzt. Der Proteintransport selbst wird durch den Exportapparat, welcher aus den fünf konservierten integralen Membranproteinen YscR/S/T/U/V besteht, und dem ATPase System ermöglicht. Zwei der Membranproteine, YscU und YscV, besitzen größere cytoplasmatische Domänen, welche eine wichtige Rolle während des Proteinexports spielen. Daher sollen in dieser Arbeit die Strukturen dieser Domänen aufgeklärt und deren Funktionen im Exportprozess untersucht werden.

Das in der inneren Membran lokalisierte Protein YscU spielt eine wichtige Rolle während des Aufbaus der äußeren Komponenten des *injectisoms*. Die cytoplasmatische Domäne von YscU ($YscU^C$) erkennt poren- und spitzenbildende Proteine (Translokatoren) als spezifisches Substrat in der Exporthierarchie. Zur Aktivierung wird YscU in einem konservierten Aminosäuremotiv der Sequenz Asn-Pro-Thr-His (NPTH) autokatalytisch gespalten. Modifikationen dieses Motivs verhindern die Spaltung von YscU, so dass die Translokatoren nicht mehr als Exportsubstrate erkannt werden und nicht mehr exportiert werden. Als Folge werden überdurchschnittlich lange Nadeln gebildet. Die Kristallstrukturen der nicht geschnittenen Varianten $YscU^C$ N263A und N263D wurden bei einer Auflösung von 2.05 Å beziehungsweise 1.55 Å gelöst. Die globuläre Domäne des Proteins besteht aus einem zentralen, gemischten β-Faltblatt, das von α-Helices umgeben wird. Das NPTH Motiv bildet einen Typ II β-*turn*, der zwei benachbarte β-Stränge miteinander

verbindet. Die NMR Analyse von geschnittenem und ungeschnittenem YscU^{C} zeigt, dass die globuläre Struktur des Proteins nach dem Schnitt erhalten bleibt. Die Struktur YscU^{C} N263D erklärt warum in YscU^{C} die Hauptkette aufgrund der optimalen Reaktionsgeometrie der Seitenkette des Asn263 über einen nukleophilen Angriff autokatalytisch gespalten werden kann. *In vivo* Analysen von YscU N263Q und H266A/R314A zeigen einen Phänotyp, bei dem der Transport von Translokatoren verhindert wird, aber normal lange Nadeln produziert. Ein Vergleich der Struktur von YscU^{C} mit verwandten Proteinen zeigt, dass die Linkerdomäne, die sich zwischen der N-terminalen Transmembrandomäne und der autokatalytischen Domäne befindet, eine gestreckte bzw. eine größtenteils α-helikale Konformation einnehmen kann. Dies erlaubt vermutlich eine optimale Platzierung der selbst schneidenden Domäne während des Exportprozesses.

Das integrale Membranprotein YscV ist die größte Komponente innerhalb des Typ III Sekretions Apparates. Ähnlich wie YscU besteht YscV aus einer N-terminalen Transmembrandomäne und einer C-terminalen cytoplasmatischen Domäne. Für das flagellare Homolog FlhA konnte gezeigt werden, dass die cytoplasmatische Domäne (FlhA^{C}) mit FlhB^{C} (YscU^{C}), der ATPase und dem ATPase Regulator interagiert. Daher wird angenommen, dass YscV^{C} zusammen mit YscU^{C} eine Ankopplungsstelle für den ATPase-Substrat Komplex bildet. Limitierte Proteolyseuntersuchungen zeigen, dass YscV^{C} aus einer Linkerregion und einer globulären Domäne besteht. Desweiteren konnte ein Reinigungsschema, das ausreichende Mengen an sauberem Protein liefert, etabliert werden. Obwohl YscV^{C} und die kürzere Variante YscV$^{C}_{356}$ unter einer Vielzahl von Bedingungen kristallisierte, war eine Röntgenstrukturanalyse aufgrund der schlechten Kristallqualität bisher nicht möglich.

Summary

The enteric pathogen *Yersinia enterocolitica* uses a sophisticated injection machinery called an injectisome or type III secretion system (T3SS) to translocate effector proteins across eukaryotic cell membranes. Upon contact with the eukaryotic host cell, e.g. a macrophage, a set of effector proteins is translocated into the macrophage's cytoplasm. There the effectors subvert macrophage cellular functions to inhibit phagocytosis of the pathogen. The supramolecular structure of the T3SS, found in many Gram-negative pathogenic or symbiotic bacteria including *Salmonella* spp., *Shigella* spp. and enteropathogenic *Escherichia coli* is evolutionarily related to the bacterial flagellum, which uses the export channel for assembling of its outer components. The protein export itself is managed by the export apparatus, consisting of five conserved integral membrane proteins YscR/S/T/U/V, and the ATPase system. Two of the integral membrane proteins, YscU and YscV, consist of large cytoplasmic domains that play an important role during the export process. Within this thesis the structures of these domains should be determined and their function within the export process should be analyzed.

The inner-membrane protein YscU plays an important role during the assembly of the outer components of the injectisome. Its cytoplasmic domain ($YscU^C$) recognizes tip- and pore-forming proteins (translocators) as individual substrates in the export hierarchy. Activation of YscU entails autocleavage at a conserved amino acid sequence Asn-Pro-Thr-His (NPTH) motif. Modification of this motif markedly changes the properties of YscU including cessation of translocator export and production of longer injectisome needles. The crystal structures of the uncleaved variants N263A and N263D of $YscU^C$ were determined at 2.05 Å and 1.55 Å resolution, respectively. The globular domain is found to consist of a central mixed β-sheet surrounded by α-helices. The NPTH motif forms a type II β-turn connecting two β-strands. NMR analysis of cleaved and uncleaved $YscU^C$ indicates that the global structure of the protein is retained in cleaved $YscU^C$. The structure of $YscU^C$ variant N263D reveals that wild-type $YscU^C$ is poised for cleavage due to an optimal reaction geometry for nucleophilic attack of the scissile bond by the side-chain of Asn263. *In vivo* analysis of N263Q and H266A/R314A YscU variants showed a phenotype that combines the absence of translocator secretion with normal needle-length control. Comparing the

structure of YscU to those of related proteins reveals that the linker domain between the N-terminal transmembrane domain and the autocleavage domain can switch from an extended to a largely α-helical conformation, allowing for optimal positioning of the autocleavage domain during the export process.

The integral membrane protein YscV is the largest component within the type III secretion apparatus. Similar to YscU, YscV consists of an N-terminal transmembrane part and a C-terminal cytoplasmic part. The cytoplasmic part of the homologues flagellar protein $FlhA^C$ has been shown to interact with $FlhB^C$ ($YscU^C$) as well as the ATPase and its regulator. Thus it has been proposed that $YscV^C$ and $YscU^C$ together form a platform for docking the ATPase-substrate complex. Limited proteolysis showed that $YscV^C$ consists of a linker region and a globular cytoplasmic domain. Furthermore, expression and a purification scheme were established yielding pure and sufficient amounts of protein. Although $YscV^C$ and its shorter variant $YscV^C_{356}$ crystallize in various conditions, as yet an X-ray analysis was not possible due to poor crystal quality.

1 Introduction

The World Health Organisation (WHO) monitors the worldwide (re)emergence of bacterial pathogens causing foodborne diseases (WHO, Fact sheet N^{o} 124). The most common pathogens inflicted in these diseases are *Yersinia enterocolitica, Salmonella spp.*, *Vibrio cholerae*, *Listeria monocytogenes* and pathogenic *E. coli.* In Germany over 200,000 cases of disease per year are reported to be caused by consumption of contaminated food (Bundesinstiut für Riskobewertung, Germany). New foodborne disease threats occur for a number of reasons. These include a steady increase in international travel and trade, pathogen resistance against antibiotics, changes in food production process, as well as human demographics and behaviour. Therefore an understanding of bacterial pathogenesis is important for developing new antibiotics and strategies against these diseases.

1.1 *Yersinia* - one of the most dangerous pathogens

The genus *Yersinia* is a coccobacillus-shaped Gram-negative bacterium, which belongs to the family of *Enterobacteriaceae*. Eleven different species are known of which three are pathogenic for humans: *Yersina pestis* and the two enteropathogenic species *Yersinia pseudotuberculosis* and *Yersinia enterocolitica* (Bergsbaken & Cookson, 2009; Wren, 2003).

Yersinia enterocolitica is one of the most common cause of foodborne gastrointestinal diseases, whereas *Yersinia pseudotuberculosis* is primarily an animal pathogen that rarely causes diseases in humans (Trulzsch *et al*., 2007). Both species are found widely in the environment and cause infection (Figure 1), after the consumption of contaminated food or water (Wren, 2003). Once the bacteria reach the small intestine, they enforce their uptake into M cells by several non-fimbrial adhesins such as invasin and the *Yersina* adhesion A (YadA) (Dersch, 2003; Isberg & Barnes, 2001; Trulzsch *et al.*, 2007). Both of these proteins interact with β 1 integrins of eukaryotic cells and are believed to mediate adherence and invasion of M cells (Isberg & Barnes, 2001). After penetrating the intestinal epithelium, the bacteria are transcytosed into Peyer´s patches and migrate to the mesenteric lymph nodes.

Subsequently, the bacteria colonize liver and spleen, where they replicate externally (Wren, 2003). As a consequence, a rapid inflammation ensues and the disease may manifest in enteritis, terminal ileitis, or mesenteric lymphadenitis with watery or sometimes bloody diarrhoea (Trulzsch *et al*., 2007). The collective form of the different symptoms is also called yersiniosis.

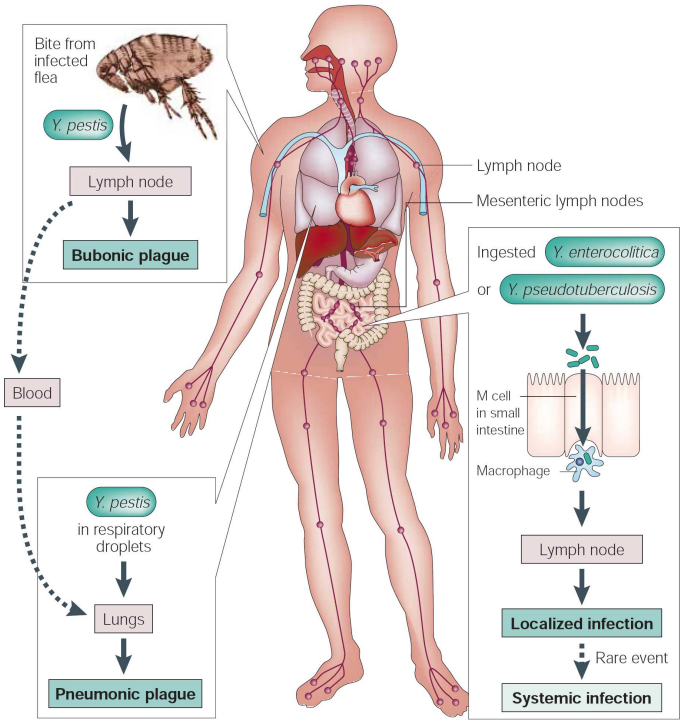


Figure 1: Path of infection of *Yersinia* spp. Plague is caused either by a flea bite or respiratory droplets from humans. In contrast, the yersiniose is caused by the uptake contaminated food or water (Wren, 2003).

The most known species *Y*. *pestis* is the causative agent of plague, which has killed millions of people during three large pandemics: the Justinian plague (5^{th}-7^{th} centuries), the mediaeval Black Death (13^{th}-15^{th} centuries) and the modern plague (1870s to present) (Perry & Fetherston, 1997). Although *Y*. *pestis* evolved from *Y*. *pseudotuberculosis* just 1,500-20,000 years ago (Achtman *et al*., 1999), it has developed a markedly different infection pathway compared to its enteropathogenic relatives (Figure 1). *Y*. *pestis* established a transmission way to non infected mammalian reservoirs (primarily rodents) using fleas as carrier. Due to additional flea plasmids, *Y*. *pestis* is able to survive in flea and is successfully transmitted during the

next bite into the mammalian host (Perry & Fetherston, 1997; Wren, 2003). Occasionally, the bacteria are injected into a human body, where they can enter the lymphatic system and cause a haemorrhaging inflammation. The resulting characteristic painfully swollen lymph nodes (bubo) are associated with this disease (bubonic plague) (Bergsbaken & Cookson, 2009; Rollins *et al*., 2003). From the lymphatic system the bacteria can enter the bloodstream, which distributes them through the whole body. In the following septicemic plague, bacterial endotoxins cause disseminated intraveascular coagulations (clots) resulting in necrosis. The infection can also progress with the bloodstream to the lung, causing the highly infectious pneumonic plague, which allows spread from person-to-person through respiratory droplets (Wren, 2003). If the plague is not treated with antibiotics on time, the mortality rate is close to 100% (Prentice & Rahalison, 2007). Due to a better hygiene and effective antibiotics, the plague is not a common health problem today.

1.2 The pVY plasmid is necessary for *Yersina´*s virulence

The first line of defence in humans against invading bacterial pathogens is the innate immune system. Bacteria possess so-called pathogen-associated molecular patterns (PAMPs) which are recognised by specific sensor proteins (pattern-recognition receptor, PRR) on the surface of immune cells. Upon detection the innate immune system activates macrophages and dendritic cells, which in turn take up the pathogens by phagocytosis and induce inflammation by releasing cytokines. Furthermore, activation of the innate immune system also set off the adaptive immune system finally to eradicate the infection (Shao, 2008). However, all three pathogenic *Yersinia* spp. seems to be largely untouched from this immune defence machinery. Genetic studies revealed that pathogenic *Yersinia* carry an extra 70 kbp plasmid, pVY, which encodes specific effectors and a type III secretion system (T3SS). The T3SS enables the bacteria to inject the effectors into the host cell and thus overcome the immune defence machinery (Cornelis *et al.*, 1998; Wren, 2003). Several other Gram-negative bacteria, like *Salmonella*, *Shigella* and enteropathogenic *Escherichia coli*, use the same nano machinery also for their infection (Cornelis, 2006).

1.3 *Yersinia* outer proteins prevent phagoytosis

During infection *Yersinia´*s T3SS delivers at least six effector proteins, which are known as Yops (*Yersina* outer proteins), into the host´s cytoplasm. These Yops work efficiently together and prevent host phagocytosis by inducing macrophage cell death. Further, the production of cytokines is suppressed, which leads to reduced inflammation and immune response (Shao, 2008; Trosky *et al*., 2008).

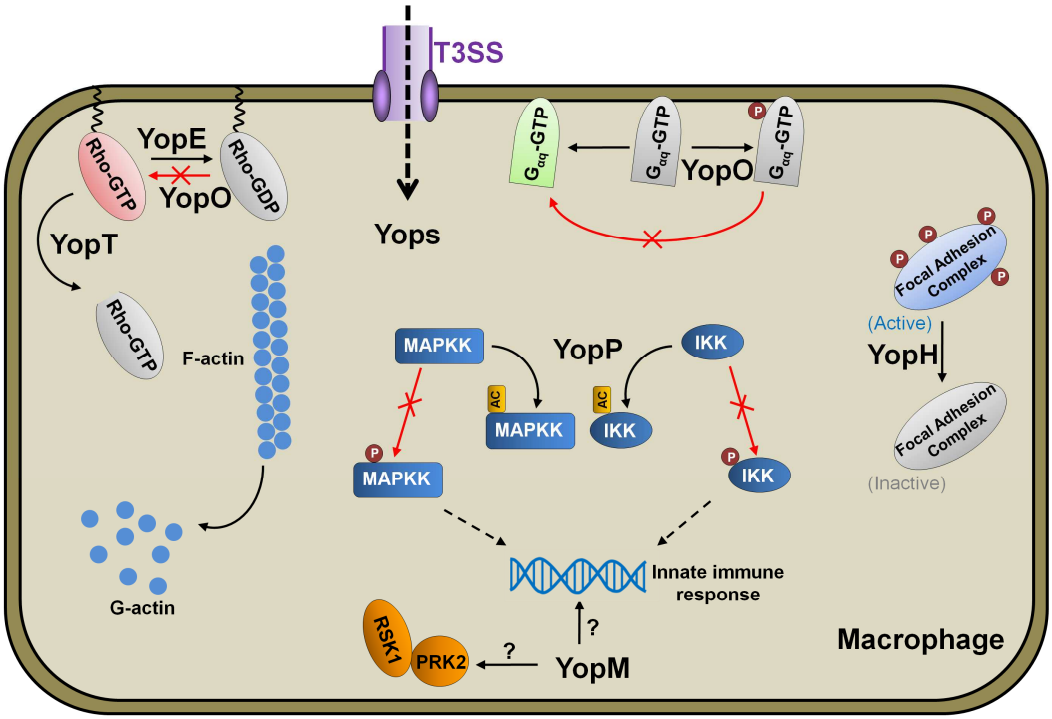


Figure 2: Targets of the six Yop effectors in the host cells. YopE, YopO, YopT and YopH interfere in the regulation of the actin cytoskeleton, which efficiently prevents phagocytosis. YopM and YopP contribute to transcriptional inhibition of immune response genes, whereas the exact role of YopM remains elusive (Shao, 2008).

The major function of YopE, YopT and YopO (YpkA in *Y. pestis* and *Y. pseudotuberulosis*) is the disassembly of actin filaments (Navarro *et al*., 2007; Von Pawel-Rammingen *et al*., 2000). They inactivate members of the family of Rho GTPases by modulating the GTP/GDP-bound states (YopE and YopO) or detach Rho GTPases from the membrane (YopT) by cleaving off the membrane anchor (Shao *et al*., 2003). YopO also blocks signalling pathways of G-proteins by phosphorylation of the Gαq subunit (Navarro *et al*., 2007). The tyrosine phosphatase YopH functions as an antiphagocytic effector by dephosphorylating several components of the focal adhesion complex and thus preventing actin polymerisation (Viboud & Bliska, 2005).

MAPK (mitogen activated protein kinase) and NF-κB (nuclear factor- κB) pathways are down-regulated by the YopP (YopJ in *Y. pestis* and *Y.pseudotuberulosis)* catalysed acetylation of MAPK kinases and IKKs (IκB kinase), which results in a suppression of cytokine production (Ruckdeschel *et al*., 2001). The role of YopM is still unclear, as effects on gene transcription revealed only contradicting results (Trosky *et al*., 2008). However, *in vitro* assays showed that the two cellular kinases protein kinase C-like 2 (PRK2) and ribosomal S6 protein kinase (RSK1) act as binding partners for YopM (McDonald *et al*., 2003). Although the kinases' activities are stimulated by binding, there is no evidence that these kinases interact during normal cellular activities (Shao, 2008).

1.4 Bacterial secretion systems

Secretion is an essential process for bacteria to modulate their environment, especially, if they interact with other organisms as symbionts or pathogens. For this purpose, bacteria have evolved a variety of molecular machineries that transport cargoes such as proteins, DNA and drugs across the bacterial membrane(s) and in some cases even across the host cell membrane. As yet, six different types of secretion systems (T1SS-T6SS) are known in Gram-negative bacteria (Figure 3). Gram-positive bacteria share some of these secretion systems and display additionally the Gram-positive specific T7SS (Tseng *et al*., 2009).

The T1SS transport proteins of diverse sizes (10-900 kDa) and functions in a single step over both bacterial membranes (Wooldridge, 2009). Thereby secretion usually depends on the presence of a specific C-terminal secretion signal. Many of the exported proteins are important for bacterial pathogenesis. The machine is made up of only three proteins: an ATPase binding cassette (ABC) transporter, an outer membrane factor (OMF) and a membrane fusion protein (MFP). T1SS also exports non-proteinaceous substrates such as cyclic β-glucans and polysaccharides (Delepelaire, 2004).

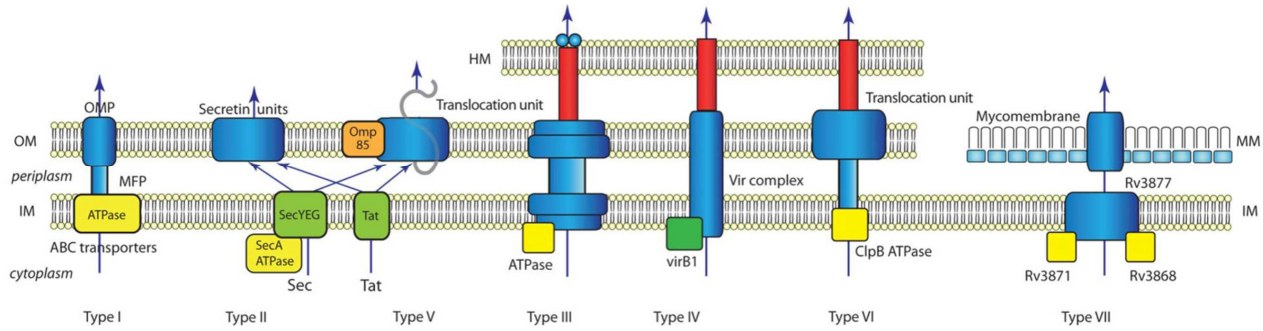


Figure 3: Schematic representation of the known bacterial secretion systems. HM, host membrane; OM, outer membrane; IM, inner membrane; MM; mycomembrane; OMP, outer membrane protein; MFP, membrane fusion protein (Tseng *et al*., 2009).

Translocation *via* T2SS and T5SS requires at least two steps. At first substrates are exported into the periplasm by the universal Sec or two-arginine (Tat) pathways and then translocated across the outer membranes. Both systems are required for export of several virulence factors. The T2SS is composed of 12 to 16 Gsp (General secretion pathway) proteins, which build a large structure (secreton) that spans the whole periplasm and is thought to connect the inner and outer membrane (Cianciotto, 2005). The T5SSs are divided into three sub-classes (T5aSS-T5cSS). All possess an outer membrane β-barrel protein domain, through which secretion of effector molecules is achieved (Henderson *et al*., 2004). The autotransporters (T5aSS) and the trimeric autotransporters (T5cSS) consist of an additional N-terminal passenger domain (40-400 kDa) that is exported into the extra-cellular space. In some case, such as the adhesion protein YadA from *Y. enterocolitica,* the passenger domain remains attached to the transporter. In other cases domains are cleaved off from the transporter and build enzymes or toxins. In the two partner secretion system (T5bSS) effector and translocator are synthesized as separate proteins that associate in the periplasm (Henderson *et al.*, 2004).

T3SS (discussed below), T4SS and T6SS translocate their substrates at least over three membranes directly into their host cells, therefore constituting virulence determinants. In comparison to the other secretion systems, the T4SS has the ability to transport DNA or protein/DNA complexes. *Agrobacterium tumefaciens* uses the machinery to inject tumor inducing plamids (Ti-plasmid) into the host cell. The resulting crown gall tumors produce opines (amine derivatives of amino acids and

sugars), which serve as a carbon source for the pathogen (Tseng *et al*., 2009). The overall architecture can be divided into a pilus and a core complex, which spans both membranes. The entire core complex (1.05 MDa) consists of multiple copies of three proteins. Recently, the crystal structure of the outer membrane channel (~0.6 MDa) has been determined at atomic resolution (Chandran *et al*., 2009). Transport energy is provided by three ATPases, which are anchored in the inner membrane (Chandran *et al*., 2009; Fronzes *et al*., 2009). The T6S machinery has been recently characterized and appears to constitute a phage-tail spike-like injectisome that induces its effectors directly into the host´s cytoplasm. A model has been proposed that includes a cytoplasmic chaperone with ATPase activity, a channel bridging both membranes and a needle that is topped by a pore structure (Shrivastava & Mande, 2008).

The T7SS is a specialized secretion system for Gram-positive bacteria like mycobacteria. In these bacteria substrates also have to be transport across the bacterial cell wall (mycomembrane), as it is heavily modified with lipids (Tseng *et al*., 2009). Structure and function of the system are still poorly understood, but a first model suggests the presence of channels in both membranes. The effector substrates are bound to chaperone-like ATPases and exported invariably as heterodimers (Abdallah *et al*., 2007).

1.5 Type III secretion system and flagellum are related

To inhibit phagocytosis, Yops have to be transported into macrophage by *Yersinia*. For this purpose, they are equipped with a T3SS, which is also called injectisome as it works like a molecular syringe. The T3SS consists of the injectisome apparatus itself, secreted proteins (translocators and effectors), chaperones and cytoplasmic regulators. About 25 Yop secretion (Ysc) proteins are required to build up this injection machinery (Cornelis, 2006). Nine of these Ysc proteins (YscC, J, N, Q, R, S, T, U, V) are highly conserved among the seven known injectisome families. Eight of them, in turn, share significant similarity with components of the flagellum, which is obvious by the overall structural similarity of T3SS and flagellum (Figure 4). Phylogenetic studies also suggest that injectisome and flagellum derive from a common ancestor (Gophna *et al*., 2003; Pallen *et al*., 2005; Troisfontaines &

Cornelis, 2005). The flagellar system uses T3S for sequential export of its outer components (rod, hook and filament). Under special conditions, the flagellar system has the ability to also export non-flagellar proteins (Young *et al*., 1999).

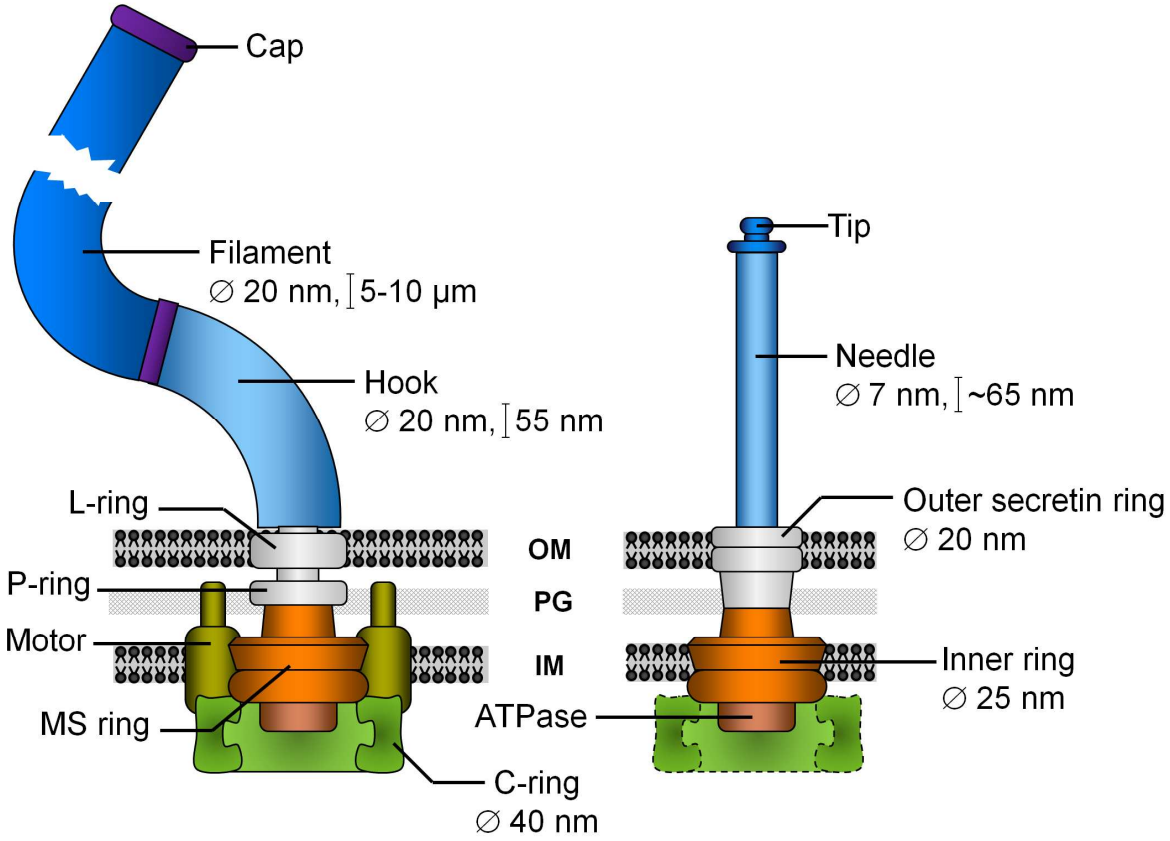


Figure 4: Schematic representation of the flagellum and the *Yersinia* injectisome. IM, inner membrane; OM, outer membrane; PG, peptidoglycan. The C-ring of the *Yersinia* injectisome is represented in dashed lines as information on this component is still scarce (Cornelis, 2006).

1.6 Architecture of the injectisome

As it is difficult to orientate in between the different T3SS proteins and their homologs, a pull-out page can be found at the end of the thesis (see appendix Figure 55). It contains a schematic drawing and sizes of the *Yersinia* type III secretion system. Further, it provides an overview on several *Yersinia* proteins and their homologues variants, which are discussed here.

Although some parts of the injectisome still remain elusive, considerable progress in structural information has been made to get insights into structure and function of the

T3SS. The molecular architecture can be roughly dived into an external part, a membrane embedded basal body, which harbors also the T3S apparatus, an ATPase complex and a proposed C-ring attached to the basal body (Cornelis, 2006) (Figure 4).

Needle, tip and pore

Depending on the injectisome the basal body is topped by a needle (Blocker *et al*., 1999; Hoiczyk & Blobel, 2001; Kubori *et al*., 1998; Marlovits *et al*., 2004), a filament (Van Gijsegem *et al*., 2000) or a pilus (Crepin *et al*., 2005; Daniell *et al*., 2001; Knutton *et al*., 1998). In *Y. enterocolitica* (strain E40, Figure 5 A and B) the injectisome is equipped with a 65 ± 10 nm long hollow stiff needle (Journet *et al*., 2003) (needle length varies from 49 ± 6 nm to 87 ± 11 nm for different strains (Wagner *et al*., 2009)), which is helically polymerised by the protein YscF. The helical parameters (5.6 units per turn, 2.4 nm helical pitch) are very similar to the flagellar hook, although the needle subunits (~ 9kDa) are significantly smaller than the hook subunits (~40 kDa) (Cordes *et al*., 2005). At the upper end the needle is terminated by a pentameric tip complex (Figure 5 A), which is made up of translocator protein LcrV. It is suggested that the tip complex serves as a scaffold or assembly platform for the translocation platform, which is made up of the translocators YopB and YopD (Mueller *et al*., 2005).

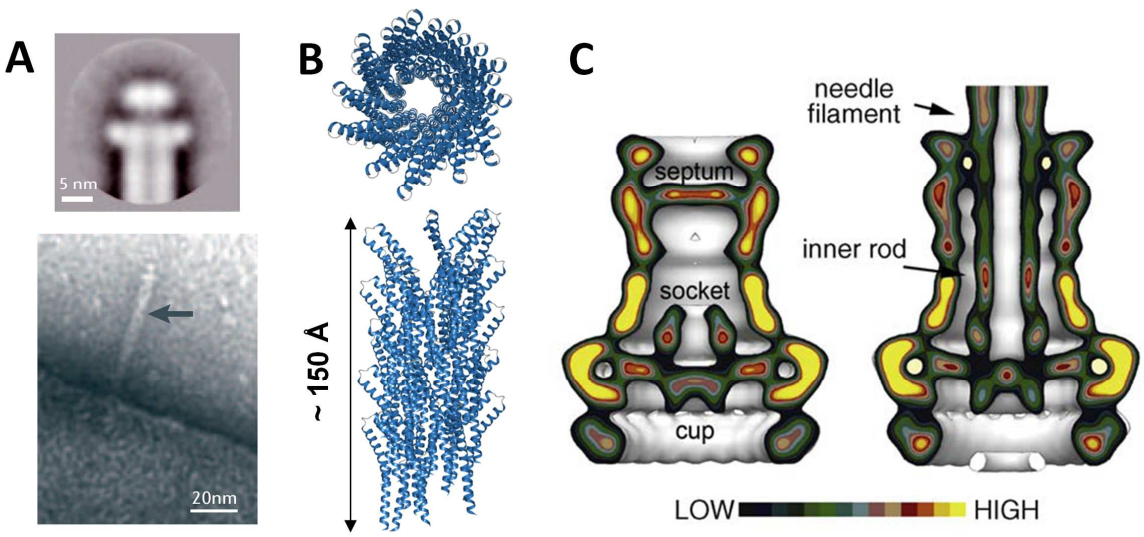


Figure 5: Structural components of the T3SS. (A) Projection average (top) of STEM images of the tip complex and transmission electron micrograph of *Y. enterocolitica* with a protruding needle (Cornelis, 2006). **(B)** Molecular model of T3S needle from *Shigella flexneri* (PDB accession number: 2V6L (Deane *et al*., 2006)). **(C)** Cutaway view of the needle complex of *Salmonella* obtained from cryo-EM (EmDep accession number 1224 (Marlovits & Stebbins, 2010)).

Basal body and C-ring

Three dimensional reconstruction of single particle cryo-EM of the needle complexes from *Salmonella* (Marlovits *et al*., 2004) and *Shigella* (Hodgkinson *et al*., 2009) provided a detailed view on the basal body and the needle base (Figure 5 C). The basal body is made up of several rings, which are embedded in the outer and inner membrane and connected by a central rod. A cutaway view of the *Salmonella* complex showed an inner channel (Ø ~ 2-3 nm) which is anchored to a socket-like structure and traverses the whole chamber of the needle complex (Marlovits & Stebbins, 2010). Analysis of the basal body revealed that the scaffold is made up of only three proteins, which are called in the *Yersinia* system YscC, D and J (Galan & Wolf-Watz, 2006; Spreter *et al*., 2009). The outer membrane rings are formed by 12 to 14 of YscC monomers that belong to the secretin family (Blocker *et al.*, 2001; Burghout *et al.*, 2004b; Koster *et al.*, 1997; Tamano *et al.*, 2000). For proper membrane insertion the assistance of lipoproteins (YscW family) is required (Burghout *et al.*, 2004a). The inner membrane rings are proposed to be built by two proteins, which belong to the conserved lipoprotein family YscJ and the less

conserved family YscD (Spreter *et al*., 2009). The crystal structure of the periplasmic domain of EscJ (Yip *et al*., 2005), the EPEC ortholog of YscJ, allowed the modelling of a 24mer inner membrane ring, whereas YscD seems to be packed onto this ring (Spreter *et al*., 2009). Within the last years, several periplasmic domain structures of all three proteins were determined and revealed common ring building motifs documenting a close evolutionary relationship between the injectisomes (Marlovits & Stebbins, 2010; Spreter *et al.*, 2009).

Adjacent to the basal body a C(ytosolic)-ring, which is made up by the proteins of the YscQ family, has been proposed (Cornelis, 2006). However, up to now a C-ring was observed only in the flagellar system, in which it is an essential component of the rotation switch (Macnab, 2003). Furthermore, an interaction between the C-ring component FliN (flagellar C-ring consists of two proteins FliM and FliN) and FliI (ATPase) -FliH (ATPase regulator) complex suggested that the C-ring serves as a platform recruiting the ATPase complex (Gonzalez-Pedrajo *et al*., 2002).

Export apparatus

In the 3D reconstructed EM picture of the *Salmonella* basal body (Figure 5) a socket-like structure, which is believed to anchor the inner rod, is found within the inner membrane ring. As it still unknown, which proteins build this structure, it is proposed to be made up of the export apparatus proteins, which build the translocation channel (Deane *et al.*, 2010). However, this hypothesis remains to be verified. The export apparatus (Figure 6) itself is made up of five highly conserved trans-membrane proteins: YscR, -S, -T, -U and -V (Cornelis, 2006). Not much known is about YscR, -S and -T so far, the cytoplasmic domains of YscU and YscV (formally known as LcrD), however, are well investigated.

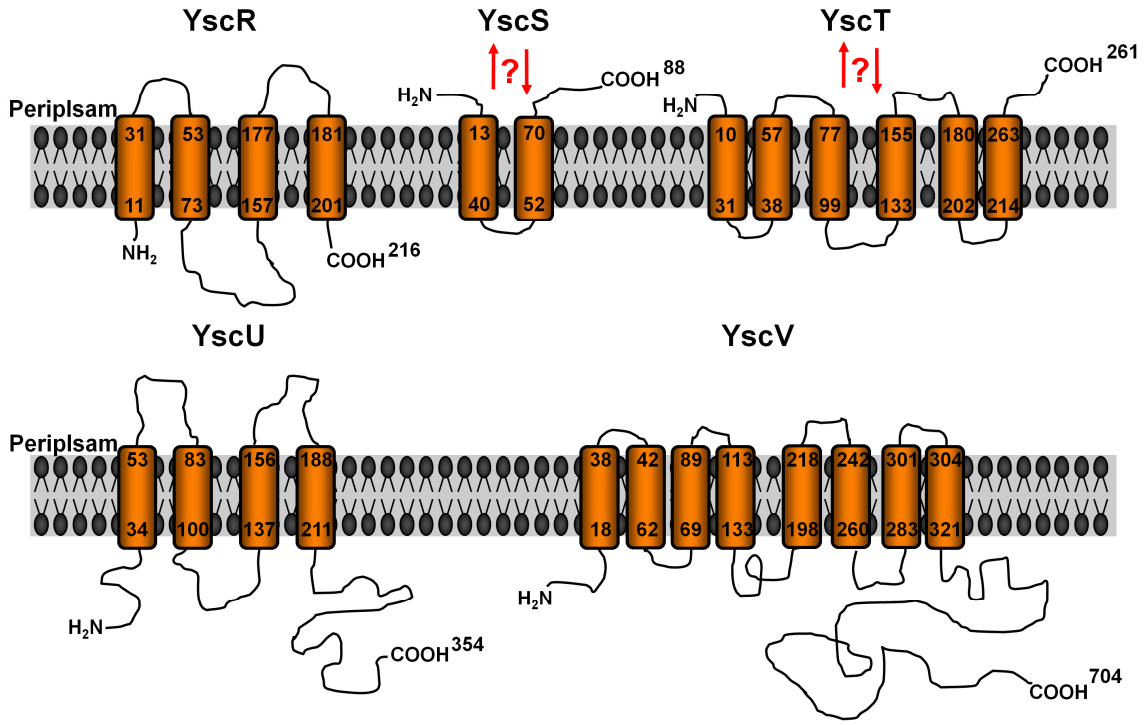


Figure 6: Proteins of the export apparatus. The orientation in the membrane of YscS and YscT still remains elusive.

In YscU the cytoplasmic domain is autocleavable (Lavander *et al*., 2002) and is part of the substrate specificity switch (Sorg *et al*., 2007) discussed below. Furthermore, it has been shown that $YscU^{C}$ interacts with YscK (Docking of ATPase to the C-ring), YscL (ATPase regulator) and YscQ (proposed C-ring) (Riordan & Schneewind, 2008). The cytoplasmic domain of FlhA, the flagellar homologue of YscV, was shown to interact with FlhB (YscU), FlhI (ATPase) and Flh (ATPase regulator) and FliJ (general chaperone) indicating an important role during T3S (McMurry *et al.*, 2004; Minamino *et al.*, 2010).

Bioenergetics and the ATPase system

For proper T3S an ATPase (YscN family) is required (Woestyn *et al*., 1994). YscN has been shown to interact with YscQ, YscK and YscL (Blaylock *et al.*, 2006; Riordan & Schneewind, 2008), but so far it is not clear how the proteins work together. EM and X-ray analyses of HrcN (ATPase of *Pseudomonas syringae*) revealed a ring structure, which is formed by two hexameric rings stacked onto each other (Pozidis *et al*., 2003). Analysis of EscN (EPEC homologue) showed that the T3SS ATPases share similarity with the F1 ATPases, but structural and sequence differences indicate a unique role during secretion (Zarivach *et al*., 2007). The ATPase InvC from *Salmonella enteric* SPI-1 has been shown to bind and dissociate chaperone-effector complexes. Further, InvC unfolds effectors by spending ATP (Akeda & Galan, 2005). The unfolding activity seems to be crucial for energizing the secretion process, but a proton motive force has been found to be equally important for the export (Wilharm *et al*., 2004). So it has been suggested that the ATPase activity is needed for both substrate recruitment and unfolding. Possibly it also supports initial transfer into the export channel, but it is believed that the proton motive force provides the energy source for the export process.

1.7 Assembly of the type III secretion system

The assembly of the injectisome can be roughly divided into two steps: the formation of the basal body including the T3S apparatus and the formation of the outer components including the needle, tip and pore. The first stage, formation of the basal body, has not been investigated so far for the T3SS, instead the flagellar system is rather very well studied (Macnab, 2004). Following the flagellar assembly scheme it is believed that basal body formation is achieved by sequential insertion of the membrane components, for example the secretin, *via* the Sec pathway (Sorg *et al*., 2007).

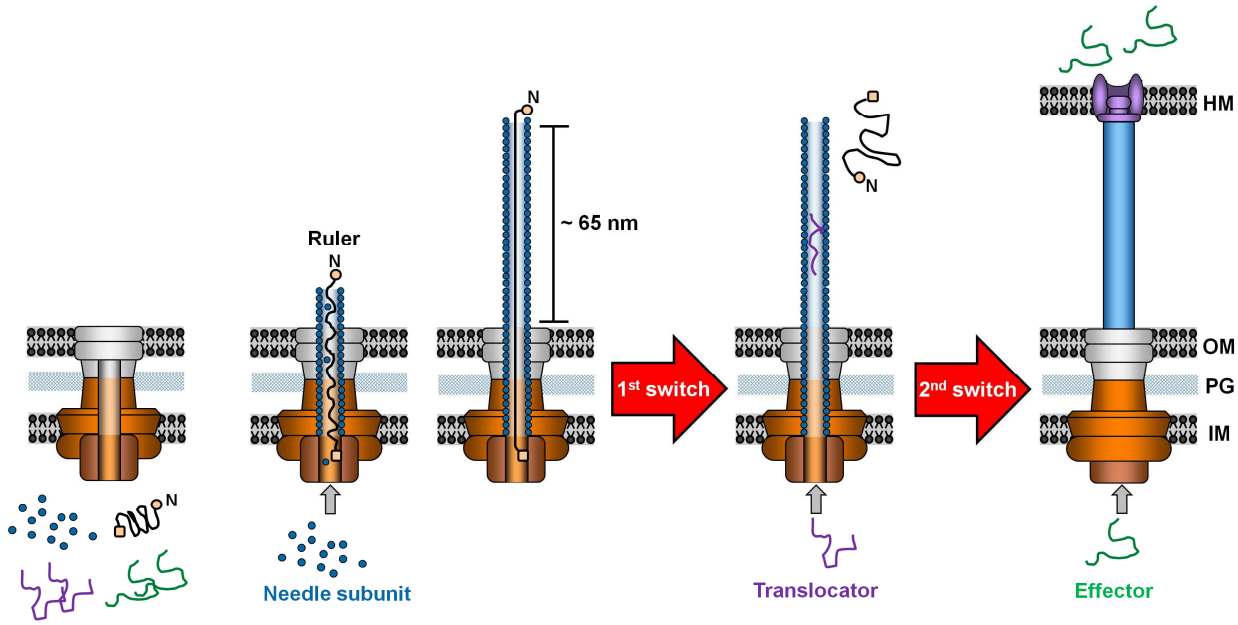


Figure 7: Assembly of the Yersinia injectisome: The ruler model. Once the basal body has been built completely, the T3S apparatus has to discriminate between different types of substrate. Early substrates such as needle subunits and the molecular ruler are transported until the needle has reached its final length. Then a switching event changes the substrate specificity to middle substrates, forming the tip and pore complex. Upon host cell contact a second switching event is triggered exporting the effectors into the host cell. HM, host cell membrane; OM, outer bacterial membrane; PG, peptido glycan layer; IM, inner bacterial membrane (Cornelis, 2006; Deane *et al.*, 2010).

Once the basal body has been inserted into the membrane, further components are transported by the T3SS in the second stage. Therefore the T3S apparatus has to discriminate between different export substrates. Needle assembly depends on the export of early substrates such as the needle subunits YscF and the length control protein YscP (Cornelis, 2006). How the needle length is controlled has been discussed intensely. Two models are currently discussed: the ruler/tape model (*Yersinia*) and the rod assembly model (*Salmonella*). The latter model suggestes that complementation of the inner rod terminates needle assembly by triggering the substrate specificity switch (Marlovits *et al.*, 2006). In the ruler/tape model, needle length is tightly controlled by the so-called molecular ruler YscP. When the needle has reached its final length, YscP triggers the substrate specificity to intermediate (translocators) substrates (Cornelis, 2006). The second protein, which is thought be involved in this switching event, is the inner membrane protein YscU, which is likely to interact directly with YscP (Cornelis, 2006; Deane *et al.*, 2010). Finally, late

substrates (effectors) are translocated into the host in a second switch event (Deane *et al.*, 2010).

1.8 The inner membrane proteins YscU and YscV

As mentioned above, the highly conserved inner membrane proteins YscU and YscV are part of the T3S apparatus and comprise 354 (~40.3 kDa) and 704 (~77.8 kDa) amino acid residues, respectively. The exact localization of the proteins is not entirely clear. They might be part of the inner ring in the basal body (Deane *et al.*, 2010). Both proteins share a similar domain architecture consisting of a N-terminal transmembrane region and a cytoplasmic domain (Saijo-Hamano *et al.*, 2005; Sorg *et al.*, 2007). YscU and YscV are anchored into the membrane with four and six transmembrane helices and the cytoplasmic domains comprise amino acid residues 211-354 (=YscU^{C}) and 322-704 (=YscV^{C}), respectively (Allaoui *et al*., 1994). The cytoplasmic domain of YscV seems to play an important role during the export process. The flagellar homologue FlhA was shown to interact with several components of the ATPase complex indicating a role in ATPase-complex recruitment. However, the mechanism of export is not clear yet.

The cytoplasmic domain of YscU, which was originally found in the C-terminal cytoplasmic domain of the *Salmonella* flagellar YscU homolog, FlhB (Minamino & Macnab, 2000a), is cleaved autoproteolytically within a highly conserved NPTH motif (Lavander *et al*., 2003). It has been shown that autocleavage is crucial for secretion in both injectisome and flagellum. Abrogated autocleavage, which can be induced by substituting either Asn or Pro by Ala, leads to a poly-hook phenotype in the flagellum (Fraser *et al*., 2003) and longer and less regulated needles in the injectisome (Sorg *et al*., 2007). Further, there is no export of translocators (YopB, YopD and LcrV), whereas effectors are not affected. This lack of translocator export is due to a lack of recognition in non cleaved YscU indicating a role of YscU not only in the substrate specificity switch (see Chapter 1.7), but also during substrate recognition (Sorg *et al*., 2007). Additionally, it was shown that cleaved GST-YscU^{C}, purified from the *Yersinia* cytoplasm, bound to YscK/L/Q, also indicating a role in recruitment of the ATPase complex (Riordan & Schneewind, 2008).

2 Aim of this thesis

In the current *Yersinia* export model, the T3SS has to discriminate between different substrate classes to assemble this complex molecular machinery. Therefore, the so-called substrate specificity switch has been proposed to involve the “molecular ruler protein” YscP and the apparatus protein YscU. For proper function the cytoplasmic domain of YscU is cleaved autoproteolytically within a conserved NPTH motif. Bacterial mutants carrying a non-cleavable mutation no longer export translocator proteins, whereas export of effector proteins is not affected. Cleaved YscU thus appears to discriminate between translocator and effector proteins. As the cleavage event and the resulting structural changes might be important for the recognition of different export substrates, the structure of $YscU^{C}$ in its cleaved and uncleaved form was the first aim of this thesis. Furthermore, based on the structures, a detailed analysis of the autocleavage should reveal its functionality.

Besides YscU, YscV also plays an essential role in T3S. Its flagellar homologue FlhA was shown to interact with several components: FlhB (YscU), FliI (ATPase), FliH (ATPase regulator) and FliJ (general chaperone). In the flagellar export model the cytoplasmic domains of FlhB and FlhA form a docking platform for the ATPase/substrate complex. Thus, the second aim of this thesis was the structural characterization of $YscV^{C}$ and biophysical analyses of potential interactions between $YscU^{C}$ and $YscV^{C}$.

3 Materials and methods

3.1 Standard materials

Molecular-biological methods used in this work are adapted from standard collections of methods and protocols (Ausubel, 2007; Coligan, 2002; Sambrook, 2000). Only variations of standard protocols are described below.

3.1.1 Chemicals, enzymes and antibodies

All chemicals were purchased from the following companies, if not stated otherwise: Difco, Eurofins MWG Operon, Fermentas, Fluka, GE-Healthcare, Hampton Research, Invitrogen, Merck, Millipore, Qiagen, Riedel de Haen, Roche, Roth, Sigma-Aldrich and Stratagene. The quality standard was "*pro analysis*" (*p.a*.).

Table 1: Enzymes and antibodies used

Enzyme	***Source***
Benzonase	Novagen
DNase I	Roche
PreScission® Protease	GE Healthcare / HZI
Trypsin	Sigma
α-Chymotrypsin	Sigma
Endoproteinase GluC	Sigma
T4 DNA Ligase	New England Biolabs
Platinum *Pfx* DNA Polymerase	Invitrogen
Pfu turbo polymerase	Stratagene
Vent DNA polymerase	New England Biolabs
Restriction endonucleases: BamHI, EcoRI, NotI	New England Biolabs
Antibody	
Anti-GST-HRP Conjugate	GE Healthcare
α-$YscU_{211-354}$ (Rabbit)	G.C. Cornelis, (University Basel, Switzerland)
α-LcrV (Rabbit)	G.C. Cornelis, (University Basel, Switzerland)

3.1.2 Molecular weight standards and kits

Table 2: Enzymes and antibodies used

Molecular weight standard	***Source***
Smart Ladder (DNA)	Eurogentec
Low molecular weight (Protein)	Fermentas
Precision Plus Protein All Blue Standards (Protein)	BIO-RAD
Kits	
QIAprep® Spin Miniprep Kit	Qiagen
QIAquick® PCR Purification Kit	Qiagen
QuikChange® Site-Directed Mutagenesis Kit	Stratagene

3.1.3 Crystallization screens

The following commercial screens (QIAGEN) were used to screen for initial crystallization conditions:

The $AmSO_4$	The PACT
The Anions	ProComplex
The Cations	JCSG+
The Classics	JCSG Core I
The Cryo	JCSG Core II
The MbClass	JCSG Core III
The MPD	JCSG Core IV
The PEGs	SM I
The pHClear	SM II

3.1.4 Oligonucleotides

Table 3: Sequences of primers used in PCR and site directed mutagenesis. Restriction- and mutation-sites are bold and underlined in the sequence. For mutagenesis only forward primers are listed

Designation	***Sequence (5'→3')***	***Restriction site***
Primers used in polymerase chain reactions		
GST-6His-YscU211-354-fwd	aaggag**ggatcc**catcaccatcaccatcacattaaggaacttaaaatgagcaagg	*BamHI*
GST-6His-YscU211-354-rev	aaggag**gcggccgc**ttataacatttcggaatgctgtttc	*NotI*

GST-6His-YscV356-704-fwd	aaggag**gaattc**catcaccatcaccatcacgcaaacaagggccgactaggg	*EcoRI*
GST-6His-YscV356-704-rev	aaggag**gcggccgc**tcataagcaaaactcgtccaagtg	*NotI*
3704	gatc**gaattc**ttataacatttcggaatg	*EcoRI*
3724	gatc**ccatgg**ccagcggagaaaagacagag	*NcoI*
Primers used for mutagenesis in E.coli		***Mutation***
$YscU^C$- N263D-fwd	catcagtggtggtagc**gat**ccgacccatattgctattg	N263D
$YscU^C$- N263Q-fwd	catcagtggtggtagct**cag**ccgacccatattgctattg	N263Q
$YscU^C$- P264A-fwd	gtggtggtagctaat**gcg**acccatattgctattgg	P264A
$YscU^C$- P264G-fwd	ctcatcagtggtggtagctaat**ggg**acccatattgctattgg	P264G
$YscU^C$- H266A-fwd	gtggtagctaatccgacc**gcg**attgctattggtattctttataagc	H266A
$YscU^C$- R296A-fwd	cccaagttcagactgtg**gcc**aaaatagcagaagaagaaggg	R296A
$YscU^C$- R314A-fwd	caacgtatcccattagcc**gct**gctctttattgggatgcgctcg	R314A
Primers used for mutagenesis in Yersinia		
4842	atcccattagcc**gct**gctctttattgg	R314A
4956	gtggtggtagct**cag**ccgacccatattg	N263Q
4958	gctaatccgacc**gct**attgctattgg	H266A
5031	gctcatcagtggtggtagctaat**ggg**acccatattgctattggtattctttataagc	P264G
5033	cgatgcccaagttcagactgtg**gcc**aaaatagcagaagaagaagg	R296A

3.1.5 Plasmids

Table 4: Plasmids used.

Plasmids	***Current strain designation***	***Genotype and derivation***	***References***
pYV plasmids			
pYV40	WT	pYV plasmid from *Y. enterocolitica* E40	Sory *et al*., 1995
pYVe22703	pYVe22703	pYV plasmid from *Y. enterocolitica* W227 serotype O:9	Cornelis *et al*., 1987
pLY4001	Δ*yscU*	pYV40 Δ*yscU;* deletion of *yscU* codons 1-354	Sorg *et al*., 2007
Expression plasmids for E. coli			
pGEX-6-P1			GE Healthcare
pGEX-6-P1_$YscU^C$	$YscU^C$	435 bp *BamH*I/*Ava*I fragment of *yscU* (211-354) in pGEX-6P-1	G.C. Cornelis, Basel, CH
pGEX-6-P1_$YscU^C$ His6	His6-$YscU^C$	453 bp *BamH*I/*Not*I fragment of *yscU* (His6 + 211-354) in pGEX-6P-1	this work
pGEX-6-P1_$YscU^C$ N263A	$YscU^C$ N263A	Site-directed mutagenesis of pGEX-6P-1_$YscU^C$	this work
pGEX-6-P1_$YscU^C$ N263D	$YscU^C$ N263D	Site-directed mutagenesis of pGEX-6P-1_$YscU^C$	this work

pGEX-6-P1_YscU^{C} N263Q	YscU^{C} N263Q	Site-directed mutagenesis of pGEX-6P-1_YscU^{C}	this work
pGEX-6-P1_YscU^{C} P264A	YscU^{C} P264A	Site-directed mutagenesis of pGEX-6P-1_YscU^{C}	this work
pGEX-6-P1_YscU^{C} P264G	YscU^{C} P264G	Site-directed mutagenesis of pGEX-6P-1_YscU^{C}	this work
pGEX-6-P1_YscU^{C} H266A	YscU^{C} H266A	Site-directed mutagenesis of pGEX-6P-1_YscU^{C}	this work
pGEX-6-P1_YscU^{C} R314A	YscU^{C} R314A	Site-directed mutagenesis of pGEX-6P-1_YscU^{C}	this work
pGEX-6-P1_YscU^{C} R296A	YscU^{C} R296A	Site-directed mutagenesis of pGEX-6P-1_YscU^{C}	this work
pGEX-6-P1_YscU^{C} H266A/R314A	YscU^{C} H266A/R314A	Site-directed mutagenesis of pGEX-6P-1_YscU^{C}	this work
pGEX-6-P1_YscV^{C}	YscV^{C}	1152 bp *EcoR*I/*Ava*I fragment of *yscV* (322-704) in pGEX-6P-1	G.C. Cornelis, Basel, CH
pGEX-6-P1_YscV^{C} 356 His6	His6-YscV^{C} 356	1071 bp *EcoR*I/*Not*I fragment of *yscU* (His6 + 322-704) in pGEX-6P-1	this work
pGEX-6-P1_YscV^{C} 356	YscV^{C} 356	1053 bp *EcoR*I/*Ava*I fragment of *yscV* (356-704) in pGEX-6P-1	G.C. Cornelis, Basel, CH
pGEX-6-P1_YscV^{C} C595S	YscV^{C} C595S	Site-directed mutagenesis of pGEX-6P-1_YscV^{C}	this work
pGEX-6-P1_YscV^{C} C703S	YscV^{C} C703S	Site-directed mutagenesis of pGEX-6P-1_YscV^{C}	this work
pGEX-6-P1_YscV^{C} C595S/C703S	YscV^{C} C595S/C703S	Site-directed mutagenesis of pGEX-6P-1_YscV^{C}	this work
Expression plasmids for Yersinia			
pBADMycHisA			Invitrogen
pLY7	*yscU*$^{+++}$	pBAD:: *yscU*; *yscU* was amplified from pYVe22703 using oligos 3704 and 3724 and cloned into the *NcoI* / *EcoRI* sites of pBADmycHisA	Sorg *et al*., 2007
pSTW7	$yscU_{N263A}{}^{+++}$	pBAD:: $yscU_{N263A}$; mutation N263A was introduced into pLY7 by site-directed mutagenesis using oligos 3725 and 3726	Sorg *et al*., 2007
pISO153	$yscU_{R314A}{}^{+++}$	pBAD:: $yscU_{R314A}$; mutation R314A was introduced into pLY7 by site-directed mutagenesis using oligos 4842 and 4843	Wiesand *et al*., 2009
pISO166	$yscU_{N263D}{}^{+++}$	pBAD:: $yscU_{N263D}$; mutation N263D was introduced into pLY7 by overlapping PCR using oligos 3724/4955 and 4954/3704	Sorg *et al*., 2007
pISO167	$yscU_{N263Q}{}^{+++}$	pBAD:: $yscU_{N263Q}$; mutation N263Q was introduced into pLY7 by overlapping PCR	Wiesand *et al*., 2009

		using oligos 3724/4957 and 4956/3704	
pISO168	$yscU_{H266A}^{+++}$	pBAD:: $yscU_{H266A}$; mutation H266A was introduced into pLY7 by overlapping PCR using oligos 3724/4959 and 4958/3704	Wiesand *et al*., 2009
pISO169	$yscU_{H266A/R314A}^{+++}$	pBAD:: $yscU_{H266A/R314A}$; mutation H266A was introduced into pISO153 by overlapping PCR using oligos 3724/4959 and 4958/3704	Wiesand *et al*., 2009
pISOA174	$yscU_{P264G}^{+++}$	pBAD:: $yscU_{P264G}$; mutation P264G was introduced into pLY7 by site-directed mutagenesis using oligos 5031 and 5032	Wiesand *et al*., 2009
pISOA175	$yscU_{R296A}^{+++}$	pBAD:: $yscU_{R296A}$; mutation R296A was introduced into pLY7 by site-directed mutagenesis using oligos 5033 and 5034	Wiesand *et al*., 2009

3.1.6 Bacterial strains

Table 5: Bacterial strains used.

E.coli strain	***Genotype***	***Source***
Top10	F⁻, *mcr* A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80 *lac*Z Δ M15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal* U *gal* K *rps*L *end*A1*nup*G	Invitrogen
XL1-Blue supercompetent cells	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´proAB lacIqZΔM15 Tn10 (Tetr)].*	Stratagene
Tuner (DE3)	F⁻ *lacY ompT gal dcm lon* $hsdS_B(r_B^- m_B^-)$ *λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	Novagen

3.1.7 Media and Buffer

Media and special buffers were subjected to autoclave sterilization (121 ℃, 2 bar, 20 min, Top7000PST, Sauter, Sulgen, Switzerland). Ampicillin (100 mg/mL) was added after media cooling below 50 ℃.

Table 6: Media and buffers used.

Designation	***Composition***
Media	
Luria Bertani (LB)	10 g/L tryptone, 5 g/L, NaCl, 5 g/L yeast extract
Terrific Broth (TB)	900 mL basic medium (12 g tryptone, 24 g yeast extract, 4 mL glycerol), completed with 100 mL buffer solution (0.17 M KH_2PO_4, 0.72 M K_2HPO_4)
SOC-Medium	20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose

Minimal Medium (MM)	Basic medium (1 g/L NH_4Cl, 3 g/L KH_2PO_4, 6 g/L Na_2HPO_4*$7H_2O$, autoclaved); completed with 100 mL supplementary solution (20% (v/v) glucose, 0.3% (w/v) $MgSO_4$, 10 mg $Fe_2(SO_4)_3$, 10 mg thiamine and 50 mg L-SeMet or 1 g $^{15}NH_4Cl$; adjust solution to pH 7.4

Buffers/Solutions

AF_4 running buffer	PBS, 0.003 M sodium azide
Blocking buffer	5% (w/v) skim milk in PBST
Cleavage buffer I	25 mM CHES (pH 7.0, 9.0 and 10.0), 100 mM NaCl
Cleavage buffer II	25 mM CHES (pH 7.0, 10.0 and 11.0), 100 mM NaCl
DNA loading buffer (10x)	10 mM Tris/HCl, pH 7.5, 0.05% (w/v) bromphenol blue, 1 mM EDTA, 50% glycerol
DTNB buffer	50 mM Tris/HCl pH 8.0, 10 mM DTNB
Epitope MBS buffer	80% (v/v) T-TBS (pH 8.0), 20% (v/v) blocking buffer (Sigma B6429), 5% (w/v) sucrose
Epitope SM-A buffer	8 M Urea, 1% (w/v) SDS, 0.5% (v/v) β-mercapto ehtanol, adjusted to pH 7.0 with acetic aicid
Epitope SM-B buffer	10% (v/v) acetic acid, 50% (v/v) ethanol, 40% (v/v) H_2O
Epitope TBS buffer	50 mM Tris/HCl, pH 7.0, 3 mM KCl, 137 mM NaCl,
Epitope T-TBS buffer	TBS, 0.05% (v/v) Tween-20
GPC buffer	20 mM HEPES pH 7.0, 150 mM NaCl
GST elution buffer	50 mM Tris/HCl pH 8.0, 150 mM NaCl, 10 mM reduced glutathione
Ion exchange buffer A	20 mM HEPES pH 7.0, 100 mM NaCl
Ion exchange buffer B	20 mM HEPES pH 7.0, 1 M NaCl
Limited proteolysis buffer	50 mM Tris/HCl, pH 8.0, 100 mM NaCl,
Lysis buffer	1xPBS, 1 tablet protease inhibitor cocktail (Complete mini, Roche), 125 U DNase (Benzonase, Merck) or spatula point DNase
NMR buffer	50 mM NaCl, 3 mM KCl, 12 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 6.8
PAGE dehydrating solution	50% (v/v) methanol, 3% (v/v) glycerol
PAGE destaining solution	40% (v/v) ethanol, 10% (v/v) acetic acid
PBS buffer	137 mM NaCl, 3 mM KCl, 12 mM Na_2HPO_4, 2 mM K_2HPO_4, pH 7.4
PBST buffer	PBS buffer, 0.1% (v/v) Tween-20
Ponceau Red staining solution	0.2% (w/v) Ponceau S, 2% (v/v) acetic acid
PreScission protease cleavage buffer	50 mM Tris/HCl, pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT
SDS-PAGE Lower buffer (4x)	1.5 M Tris/HCl, pH 8.8
SDS-PAGE running buffer	25 mM Tris/HCl, 192 mM glycine, 0.1% (w/v) SDS
SDS-PAGE sample buffer (2x)	1 mL 1 M Tris/HCl pH 6.8, 2.4 mM glycerol, 0.8 g SDS, 2 mg Coomassie blue G-250, 0.31 g DTT, add H_2O to 10 mL
SDS-PAGE separating gel	10 mL Polyacrylamide solution, 5 mL 4 x Lower buffer, 0.2 mL SDS (10%), 4.7 mL H_2O, 20 µL TEMED, 50 µL APS (25%)
SDS-PAGE stacking gel	1.5 mL Polyacrylamide solution, 2.5 mL 4 x Upper buffer, 5.9 mL H_2O, 15 µL TEMED, 25 µL APS (25%)
SDS-PAGE staining solution	30% (v/v) ethanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie R-250
SDS-PAGE Upper buffer (4x)	0.5 M Tris/HCl pH 6.8, 0.4% (w/v) SDS
TAE buffer	40 mM Tris/HCl pH 8.2, 20 mM sodium acetat, 1 mM EDTA
Transfer buffer	20 mM Tris pH 8.0, 192 mM glycine, 15% methanol

3.2 Molecular cloning

Bacterial strains, plasmids and genetic constructions are given in Table 4 and Table 5 *E. coli* Top10 or XL1-Blue strains were used for plasmid purification and cloning. Bacteria were grown routinely on Luria-Bertani (LB) agar plates and in liquid LB-medium. Ampicillin was used at concentrations of 200 µg/mL.

Plasmids were generated using either the *Pfu* turbo polymerase (Stratagene) or Vent DNA polymerase (New England, Biolabs). Oligonucleotides for genetic constructs are given in Table 3. All constructs were confirmed by sequencing using a 3100-Avant genetic analyser (ABI Prism) or by the companies GATC (Konstanz, Germany) or MWG (Ebersberg, Germany).

3.3 Protein production and purification

3.3.1 Expression and purification of YscU^{C}

Standard expression of YscU^{C}

Testing different media, temperatures and *E. coli* cells, the following optimal expression conditions were found. The YscU^{C}-pGEX-6P-1 vector was transformed into *E. coli* Tuner (DE3) cells and cultivated on LBA agar plates overnight. For further preparations and long term storage a glycerol culture was produced. A single colony was used to inoculate a pre-culture of 20 mL LBA medium. 10 mL pre-culture, which had been grown at 37 °C and 180 rpm to an OD_{600} of 0.6, was transferred into 1 L main-culture (TBA medium). After the main culture reached an OD_{600} of 0.6 (37 °C, 110 rpm), expression of *yscU^{C}* was induced by adding 0.2 mM isopropyl thiogalactopyranoside. Shifting the temperature to 20 °C prevented the formation of inclusion bodies. After incubation for further 16 h, cells were harvested by centrifugation (6,080 xg, 15 min, 4 °C), resuspended in ice cold lysis buffer and stored at -20 °C for future preparations.

Production of labelled protein of YscU^{C}

L-Selenomethionine (SeMet)-labelled and ^{15}N-labelled YscU^{C} N263A, as well as ^{15}N-labelled YscU^{C} wild type variants were produced by using a protocol based mainly on (Guerrero *et al*., 2001). All variants were freshly transformed and plated on LBA agar plates. A single colony was used to inoculate an overnight culture of LBA medium. Cells were centrifuged (2,700 xg, 4 °C, 10 min), washed carefully in ice cold sterile ddH_2O and again centrifuged (2,700 xg, 4 °C, 10 min). The bacterial pellet was diluted in the desired minimal medium to an OD_{600} 1.0 and incubated for 20 min at 37 °C. Expression was induced by adding 0.2 mM isopropyl thiogalactopyranoside and further incubation at 20 °C for 20 h. Finally cells were centrifuged (6,080 xg, 4 °C, 15 min) and resuspended in lysis buffer.

Purification of YscU^C

Efficient cell lysis was performed by running lysis suspension twice through a homogenizer (20,000 kPsi, 4 °C). Cell debris and in soluble components were separated by centrifugation (37,060 xg, 45 min, 4 °C). Supernatant was incubated with equilibrated (PBS buffer) glutathione sepharose for 4 h at 4 °C. Unbound protein was washed three times (50 mL) from the GS column with protease buffer.

To obtain GST-labelled YscU^C, sepharose was incubated five times (50 mL) with GST elution buffer. GST-YscU^C wash and elution fractions were analysed by SDS-PAGE and concentrated (10 - 15 mg/mL).

The GST fusion tag was cleaved off by adding 75 µL PreScission protease (500 U) in 10 mL protease buffer and incubation (4 °C) overnig ht. Only YscU^C is eluted from the column, as the protease is also GST-tagged. Finally the GST-sepharose was washed with 10 mM reduced glutathione in 10 mL protease buffer to control cleavage efficiency. Purity and cleavage efficiency were analyzed using a SDS-gel. YscU^C containing fractions were pooled and dialysed against IEC buffer A (SnakeSkin membrane, MWCO 3,500, Thermo Scientific).

As YscU^C is positively charged at pH 7.0 (theoretical pI 8.1, calculated with Vector NTI InforMax), the second purification step was performed using cation exchange chromatography (MonoS 10/100 column). Runs were programmed containing two linear step gradients: 1st 0-30 % IEC buffer B in 3 column volumes and 2nd 30-100% IEC buffer B in 1 column volume. YscU^C was eluted as a single peak at ~19 % IEC buffer B (~ 20 mS/cm). YscU^C containing fractions were analyzed by SDS-PAGE, pooled and concentrated to 5 – 15 mg/mL (Vivaspin concentrator, MWCO 10,000). As final purification step gel permeation chromatography was performed using a Superdex 16/60 prep grade column (GE Healthcare).

3.3.2 Expression and purification of YscV^C

Expression of YscV^C

After transformation of YscV^C-pGEX-6P-1 into *E.coli* Tuner (DE3) cells, over-expression was performed as described for YscU^C in Chapter 3.3.1. For long term storage and further preparations glycerol cultures were produced.

Purification

Resuspended cells were lysed by running cell suspension twice through a homogenizer (20,000 kPsi, 4 °C). Cell debris and in soluble components were separated by centrifugation (37,060 xg, 45 min, 4 °C).The supernatant was incubated

with 10 mL glutathione sepharose (equilibrated in PBS buffer) for 4 h at 4 ℃. Unbound material was washed three times (50 mL) from the GS column with PreScission protease cleavage buffer sepharose was further incubated four times (3 x 10 mL, 1 x 20 mL) with GST elution buffer and analyzed by SDS-PAGE. GST-$YscV^C$ fractions were pooled and a 75 µL sample of PreScission Protease (GE Healthcare, 500 U) was added to cleave the GST fusion protein. After 45-60 min of protein cleavage, small amount of protein precipitation was observed and cleavage was immediately terminated by adding one tablet of Complete Mini (Roche). In a second step the protein was further purified using an ion-exchange column. As $YscV^C$ is negatively charged at pH 7.0 (theoretical pI 5.2, calculated with Vector NTI InfroMax), protein solution was applied to a MonoQ 10/100 column (GE Healthcare). The column was equilibrated with ion exchange buffer A and $YscV^C$ was eluted in a single peak within a linear gradient (0.1 M -1 M NaCl) at 35 % ion exchange buffer B. $YscV^C$-containing fractions were identified by SDS-PAGE and concentrated to 10 – 15 mg/mL.

3.4 Protein analytical methods

3.4.1 Discontinuous SDS-polyacrylamide gel electrophoresis

A discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was carried out to analyse the purity, amount and molecular weight of the proteins (Laemmli, 1970). The SDS-gels used in this work were performed with a 5 % (w/v) acrylamide stacking gel and a 15 % (w/v) acrylamide separating gel (Chapter 3.1.7). Protein samples (2-15 µL) were incorporated with SDS-PAGE sample buffer (10-20 µL) and briefly (60 sec) boiled at 95 ℃. 12 µL of these samples and 5 µL marker were loaded on a SDS-gel. The SDS-PAGE was run constantly at 40 mA and terminated upon reaching the dye front at the end of the gel. After electrophoresis, gels were stained in a Coomassie Brilliant Blue R-250 solution by boiling and shaking. To remove excess stain, gels were destained. Finally gels were dried between cellophane papers for storage.

3.4.2 Limited proteolysis

Limited proteolysis was used to identify globular and stable protein domains. Therefore purified protein was diluted in limited proteolysis buffer to a final concentration of 0.3 – 1 mg/mL in a 100 µL reaction volume. typsin, α-chymotrypsin and endoproteinase GluC were added in a protein/protease ratio of 1:200 (w/w) in each trial reaction. After distinct time points 10 µL aliquots were taken and proteolysis was stopped by mixing with 10 µL SDS-PAGE sample buffer (0.1 mM AEBSF (4-(2-aminoethyl)- benzenesulfonyl fluoride)) and heating for 60 sec at 95 ℃. Until further analysis by SDS-PAGE all samples were frozen at -20 ℃.

3.4.3 pH-dependent cleavage of YscU^C variants

The different YscU^C variants were expressed and purified as described above. Variant P264A was dialysed against 25 mM Ches-buffer, 100 mM NaCl for 3h at room temperature. P624G, H266A/R314A and N263D were dialysed against 25 mM CAPS-buffer, 100 mM NaCl for 4h at room temperature. Afterwards, the variant P264G was stored for 48 h and H266A/R314A and N263D were stored for 24 h at 4 ℃. Cleaved proteins were submitted to SDS-PAGE (15%) and stained with Coomassie brilliant blue.

3.4.4 Western blot and immunodetection

Protein samples from SDS-gels were transferred *via* electrophoresis on a polyvinyl difluoride (PVDF) membrane (Towbin *et al*., 1979) and used for immunodetection or N-terminal sequencing. The non-stained SDS-gel was incubated for 15 min in transfer buffer and placed onto the activated PVDF membrane (incubation 60 sec in methanol and 5 min in transfer buffer). Membrane and gel were placed in between two transfer-buffer-soaked Whatmann tissues and fixed in the semi-dry blot chamber (TransBlot semi-dry transfer cell, Bio-Rad). Transfer was carried out at constant 15 V for 45 min and successfully transfer was evaluated by Ponceau Red staining.

For immunodetection the blotted PVDF membrane was incubated for 1 h in 50 mL blocking buffer and washed three times with PBST buffer (10 mL, 5 min). Primary antibodies were diluted (α-GST-HRP conjugate 1:5000, α-YscU^C 1:1000) in PBST buffer and incubated with PVDF membrane for 1 h and again washed three times with PBST buffer. A secondary α-rabbit-HRP-antibody (1:5000) for α-YscU^C was necessary for detection and therefore the membrane was incubated and washed a second time. To detect target proteins the Enhanced Chemiluminescence kit (ECL) (LumiLight, Roche) was used.

3.4.5 Qualitative determination of accessible Sulfhydryl groups

The accessibility of protein sulfhydryl groups was determined qualitatively using Ellman`s regant (5,5'-Dithio-bis(2-nitrobenzoic acid, DTNB). 2 µL DTNB-solution (10 mM in 0.1 M Tris pH 8.0) were incubated with 1 µL protein solution (1 mg/mL) for 5min. The release of TNB (2-nitro-5-mercaptobenzoic acid) during the reaction was measured photometrically at 412 nm.

3.4.6 Mass spectroscopy

For protein identification and conformation of correct protein labelling samples were analysed using MALDI-TOF-MS (Matrix Assisted Laser Desorption Ionisation – Time Of Flight) or ESI-MS (Electrospray Ionisation). The experiments were carried out by Dr. Manfred Nimtz and Undine Felgenträger (HZI, Braunschweig).

3.4.7 N-terminal sequencing

For evaluating and identification of proteins the N-termini were sequenced by automated Edman degradation (Edman & Begg, 1967). The protein samples were separated by SDS-PAGE and blotted onto PVDF membranes as described in Chapters 3.4.1 and 3.4.4. Blotted proteins were stained with Ponceau red and the corresponding bands were cut out for sequencing. N-terminal sequencing was performed by Rita Getzlaff (HZI, Braunschweig) using the 494A HT Protein Sequenzer (Applied Biosystems).

3.4.8 ThermoFluor assay

The ThermoFluor assay is a rapid screening method to identify optimal protein stabilizing conditions by varying pH, small molecules and buffer. A fluorescent dye, which has an affinity for the hydrophobic core of proteins, is added to the protein solution. In the following a differential fluorimetry scan is performed creating a temperature gradient. Once the melting temperature of the protein is reached, fluorescence is increased due to unfolding of protein. A simple calculation affords the exact melting temperature. The highest given melting temperature corresponds to the best stabilization conditions (Niesen *et al*., 2007).

Differential scanning fluorimetry was performed on a CFX96 Real-Time Thermal Cycler C1000 (Bio Rad) instrument. In order to choose appropriate scanning conditions an initial scan was carried out varying protein and dye concentrations. For this 40 µL protein buffer were mixed with 5 µL protein (0.1, 1 and 2 mg/mL) and Sypro Orange 5000x (10-, 50- and 100x) in a 96-well plate (Multiplate 96-Well Unskirted PCR plate, Bio Rad). The plate was sealed using a non-fluorescence absorbing seal (Microseal 'B'. Adhesive Seal, Bio Rad).The best protein/dye combination (increase of fluorescence > 800 U) was chosen for the main screen, in which the different buffer conditions were tested. The conditions were adapted from a protocol of Vedadi group (Niesen *et al*., 2007). Runs were performed using a linear gradient (10 - 90°C) measuring fluorescence intensity every 0.5 °C. The data were analyzed using the CFX software (Version 1.5, Bio Rad).

3.4.9 Epitope scan

An epitope scan is a method for investigating interaction sites of proteins and also for determining the influence of amino acids on the binding affinity (Dooley & Houghten, 1993). For this purpose one binding partner is split into peptide fragments. Theses peptides were synthesized and spotted on a cellulose membrane. The potential binding partner is incubated with the membrane and binding is analysed by immunodetection (Figure 8).

The choice of the correct peptides depends on previous information about the two potential binding partners. If there is no information on binding behaviour, one protein is fragmented into overlapping peptides and incubated with its potential partner. The standard peptide length is a 15mer, which will be shifted by three amino acids along the primary protein sequence. Such a rough epitope scan is called a 15/3 scan and is used for localisation of binding sites (Dooley & Houghten, 1993). For exact localisation a 15/1 run is used afterwards. Once the binding areas are detected, the role of individual amino acids can be investigated. In these further scans each amino acid in the peptide is exchanged sequentially against either alanine (alanine scan) or every other amino acid (amino acid scan) (Frank, 2002).

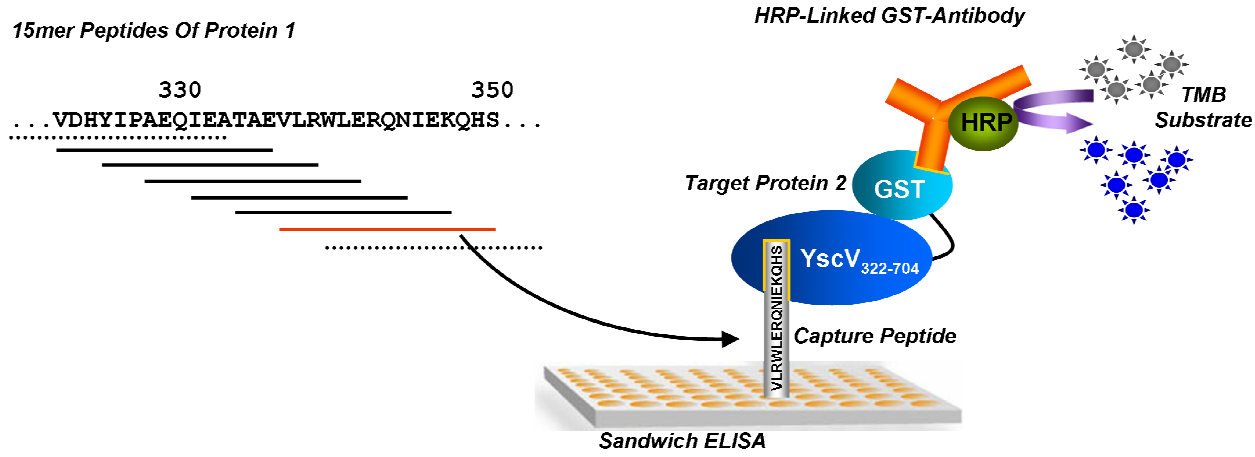


Figure 8: Functionality of the epitope scan. A GST-tagged target protein is captured by an anchored peptide and detected via an HRP-linked GST-antibody.

All peptides were synthesized and spotted onto cellulose membranes by Dr. Ronald Frank and Dr. Werner Tegge (HZI, Braunschweig). The prepared membrane was soaked in ethanol, washed three times with TBS buffer (10 mL, 10 min) and incubated in 10 mL MBS buffer at room temperature overnight. After washing (10 mL T-TBS, 10 min) once, the membrane was incubated with a 1 - 4 mg/mL protein solution diluted 1:200 in 10 mL MBS buffer for 4 h. Again the membrane was washed three times (10 mL T-TBS, 10 min) and incubated with a GST-HRP conjugate diluted 1:5000 in 10 mL MBS buffer for 2 h. After a further washing round (three times 10 mL T-TBS, 10 min), the membrane was stained using the Enhanced Chemiluminescence kit (ECL) (LumiLight, Roche).

For reuse of the membranes proteins and antibodies were removed in the following way: The membrane was washed in 20 mL water (three times for 10 min), 20 mL SM-A (three times for 10 min), 20 mL SM-B (three times for 10 min) and finally in 20 mL EtOH. Membranes were air-dried and stored in a shrink-wrap film at -20 ℃ for future use.

To test the membranes for unspecific interactions all membranes were initially incubated with GST and antibody. Control and experiment were repeated two times, respectively.

3.4.10 Dynamic light scattering

All purified protein samples were analyzed by dynamic light scattering (DLS) to assess their polydispersity. The method was also used to confirm the estimated molecular weight assuming a globular shape. Since huge particles disturb measurements, protein solutions (0.5 – 1 mg/mL) were centrifuged (20 min, 14,000 xg) and 15 µL sample volumes were transferred in a dust-free DLS-cuvette. The experiments were carried out at room temperature on a DynaPro Titan Temperature Controlled MicroSampler instrument (Wyatt Technologies). The laser power was set to yield at least 1,000,000 counts/sec. To analyze data, DYNAMICS V6 (Version 6.10.1.2) software was used. Protein solutions with suitable polydispersity (< 20%) were used for crystallization trials.

3.4.11 Asymmetric flow field flow fractionation

Asymmetric flow field flow fractionation (AF4) is a one-phase chromatography technique (Figure 9), which enables a sharp and gentle separation of protein solutions in flow fields (Giddings, 1993). In combination with static light scattering and refraction measurement the received fractions can be analysed towards oligomerisation status and concentration. The separation process needs three steps. During injection (step 1) the sample is collected under the injection port due to a balanced flow. At the same time the sample is focussed (step 2) in a thin band towards the membrane. Diffusion associated with Brownian motion, in turn, creates a counteracting motion. Smaller proteins, which have a higher diffusion rate, reach upper levels and thus a molecular weight gradient is built. To separate and elute the proteins a second perpendicular flow is injected (step 3). Because of the laminar flow, the flow field is parabolic and thus smaller molecules reach a higher velocity. In contrast to gel permeation chromatography the smaller proteins in AF4 will be eluted first (Podzimek *et al*., 2008).

All experiments were carried out on an Eclipse 3 AF4 system (Wyatt Technology), which was combined with a Dawn Heleos II Lightscattering K5 MALS device (multiple angle light scattering, Wyatt Technology) and Optilab rEX refractometer (Wyatt Technology). As the technique is sensitive to dust the running buffer was filtered (0.2 µm) twice. Sodium azide (0.003 M) was added to prevent fungal and bacterial growth. Runs were performed in a small channel (size S) equipped with a 0.35 µm spacer and 10 kDa cellulose membrane. The detector flow and the focus flow were set to 1.0 ml/min and 1.5 mL/min, respectively. Samples were injected with a flow rate of 0.2 ml/min. Channel separation quality was checked by using BSA before and

after each run. Data were processed with the ASTRA software (Version 5.3.2.20, Wyatt Technology Corporation).

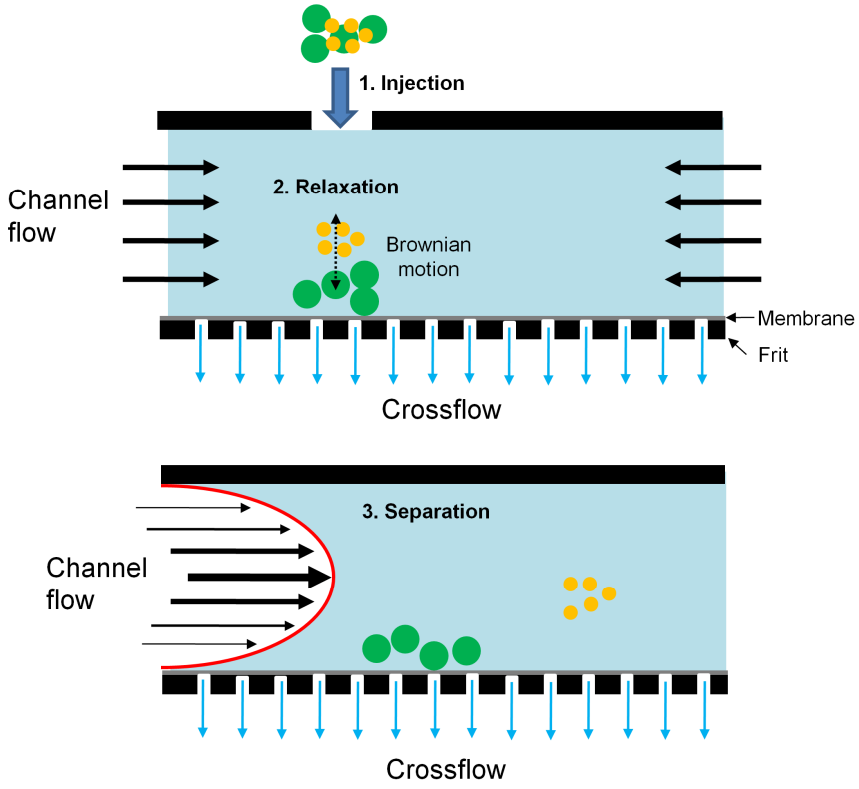


Figure 9: Principle of the AF4 technique. At first the sample is injected into a channel, where it focused and equilibrated in a thin layer. Due to a further perpendicular and parabolic flow, particles were separated along the channel.

3.5 *In vivo* analyses

Within the scope of the injectisome collaboration between the HZI and the group of Guy R. Cornelis (University of Basel, Switzerland) all *in vivo* experiments with *Yersinia enterocolitica* were carried out by Dr. Isabel Sorg, Marlise Amstutz and Stefanie Wagner. The aim was to characterize different YscU variants with respect to their secretion behaviour and their needle length.

3.5.1 LcrV secretion and immunoblotting

Induction of the *yop* regulon was performed as described by Cornelis *et al.*, 1987. Expression of the different genes cloned downstream of the pBAD promoter was induced routinely by adding 0.2% L-arabinose to the culture before shift from 28 °C to

37 ℃, and again 2 h later. Glycerol (4 mg/ml) was added as a carbon source when expressing genes from the pBAD promoter. Total cell and supernatant fractions were separated by centrifugation at 20,800g for 10 min at 4 ℃. The cell pellet was taken as the total cell fraction. Proteins in the supernatant were precipitated with 10% (w/v) trichloroacetic acid (final concentration) for 1 h at 4 ℃.

Supernatant (SN) and total cell (TC) fractions were separated by 12% SDS-PAGE. In each case, proteins secreted (SN) or produced (TC) by 2.5 x 10^7 bacteria were loaded in each lane. Immunoblotting was performed using rabbit polyclonal antibodies against LcrV (MIPA220; 1:2000). Detection was performed with the respective secondary antibodies conjugated to horseradish peroxidase (1:5000; Dako), before development with supersignal chemiluminescent substrate (Pierce).

3.5.2 Electron microscopy

T3SS needles were visualized by transmission electron microscopy as described (Agrain *et al.*, 2005; Hoiczyk & Blobel, 2001). After induction of the *yop* regulon for 4 h at 37 ℃ bacteria were harvested by centrifugatio n at 2000 xg and resuspended gently in 20 mM Tris-HCl, pH 7.5. Droplets were applied for 1 min to freshly glow-discharged, formvar-carbon coated grids, and negatively stained with 2 % (w/v) uranyl acetate. Bacteria were visualized in a Philips CM100 transmission electron microscope at a nominal magnification of 20,000 xg and an acceleration voltage of 80 kV. Needle sizes were measured with the Soft imaging system software (Hamburg, Germany).

3.6 Protein crystallization

Concentrated and purified proteins were filtered (0.2 µm) or centrifuged (20 min, 18,800 xg, 4 ℃) before crystallizing. All proteins were crystallized by hanging- and sitting-drop vapour diffusion methods at constant temperatures (20 or 4 ℃).

Initial screening

Initial screening experiments for crystallization conditions were performed with commercially available screens (Quiagen, see Chapter 3.1.3). For this purpose the sitting-drop method in 96-well plates (Greiner) was used for crystallization. 300 nL purified protein and an equal amount of reservoir were mixed in a droplet using a pipetting roboter (Mosquito robot, TTP LabTech). Plates were sealed using a Manco™ Crystal Clear tape (Jena Bioscience) and incubated at 4 or 20 ℃.

Optimization

Initial hit conditions were optimized manually in 24-well hanging drop and sitting drop formats. Drop volumes were enlarged to 2-5 µL and physico-chemical parameters such as precipitant, protein concentration, pH, ionic strength and temperature were varied producing crystals suitable for X-ray diffraction experiments.

Macro- and micro-seeding (Bergfors, 2003; Luft & DeTitta, 1997), Additive Screen (Hampton Research), varying crystallization temperatures, truncation of flexible regions (Dale *et al*., 2003) and cysteine replacements were used for further crystal improvement and to get larger crystals.

Cryo-crystalography

As X-ray radiation causes radiation damage in protein crystals (Garman & Owen, 2006), data sets were collected at cryogenic temperatures (~100 K) to minimize these effects. Upon freezing, however, water can also crystallize and water ice imposes a mechanical stress on the protein crystal resulting in loss of diffraction or ice-rings on the diffraction pattern. Therefore a cryo-protection for protein crystals is needed. This is achieved by adding cryo-protectants such as PEG 400, glucose, ethylene glycol or glycerol. In this work an aliquot of the original reservoir solution was mixed with glycerol to yield a final concentration of 15-20% (v/v), depending on the actual crystallization conditions. Crystals were harvested with a nylon loop, quickly transferred into the cryo-protecting solution and then flash frozen in liquid nitrogen.

3.7 X-ray structure analysis

Single crystal X-ray crystallography is a powerful method to determine the three dimensional structures of proteins at atomic resolution. Protein crystals are irradiated with a monochromatic X-ray beam. Upon interaction of X-ray radiation and matter (protein crystal), electrons interact with the changing electromagnetic field and in turn begin to oscillate. As oscillating dipoles, they emit coherent radiation with the same wavelength as the incident radiation. This effect is called elastic or Thomson scattering. Under certain crystal orientations the emitted waves interfere constructively (Bragg's law), yielding detectable reflection spots. The symmetry of the resulting reflection pattern directly corresponds to the crystal lattice, while the intensity distribution depends on the electron density distribution within the unit cell. To calculate the electron density (Equation 1) the phase information is required, which is lost during data collection. This problem is known as the phase problem, which can be solved by a variety of indirect methods.

In this study the multi-wavelength anomalous dispersion (MAD) technique was used to solve the phase problem. For this purpose a selenomethionine (SeMet) derivative protein has been produced (see Chapter 3.3.1) for structure determination (Hendrickson *et al*., 1990).

$$\rho(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l |F_{hkl}| \cdot e^{-2\pi i(hx+ky+kz-\alpha'_{hkl})}$$ **Equation 1**

With phase $\alpha = 2\pi\alpha'$

3.7.1 Phase determination *via* anomalous diffraction

If incident X-rays are reemitted by atoms with the same phase angle, then an elastic scattering is observed. But when the X-ray wavelength is in the range of the scattered electron binding energies (heavy-atom absorption edge), a part of the radiation is absorbed by the heavy atoms. As a consequence of this inelastic interaction (Compton scattering) the reemitted intensity decreases and the phase is altered compared to the incident wave. This effect is designated as anomalous diffraction and can be accounted for the atomic scattering factor f as two perpendicular wavelength-depended terms (Equation 2): a real term f' (dispersive) and an imaginary term f'' (absorption). More importantly Friedel´s law ($|\underline{F}_{hkl}|=|\underline{F}_{-h-k-l}|$) loses its validity. Due to the wavelength-depended contributions of the anomalous scatterer, the phase as well the amplitude of symmetry related reflections changes and thus Friedel mates differ (Equation 3). These differences of Friedel mates are used to obtain phase information and substructure handedness in the MAD experiment (Figure 10).

$$f = f_0 + f'(\lambda) + if''(\lambda)$$ **Equation 2**

f_0	wavelength independent atomic scattering factor
$f'(\lambda)$, $if''(\lambda)$	real and imaginary atomic components of anomalous scattering

$$\underline{F}(+) = \underline{F}_{PH} + \underline{F'}(+) + \underline{F''}(+)$$
$$\underline{F}(-) = \underline{F}_{PH} + \underline{F'}(-) - \underline{F''}(-)$$ **Equation 3**

$\underline{F}(+)$, $\underline{F}(-)$	structure factors of Friedel mates
$\underline{F}_{PH}$	structure factor of a heavy atom derivative at normal wavelength
$\underline{F'}$, $\underline{F''}$	reel and imaginary components of anomalous scattering

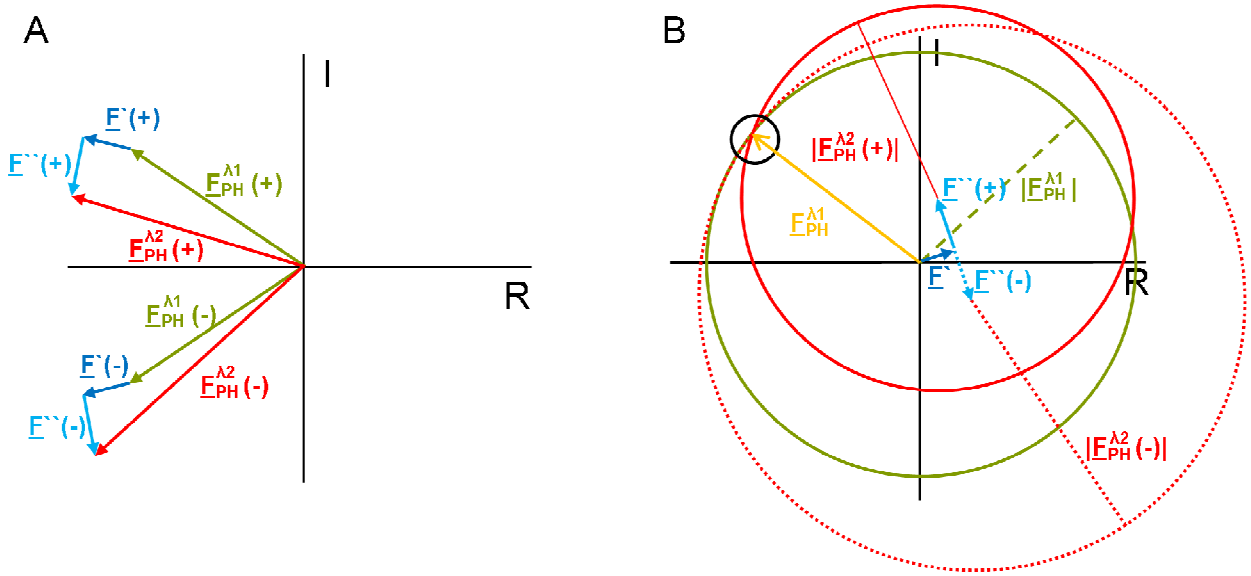


Figure 10: (A) Construction of the Friedel mates at anomalous scattering. Under anomalous scattering, at wavelength λ2, $\underline{F}_{PH}^{\lambda 2}(+)$ is no longer the mirror image of $\underline{F}_{PH}^{\lambda 2}(-)$.

(B) Graphical vector solution using a Harker construction for MAD phasing. Once the structure factor $\underline{F}_{PH}^{\lambda 1}$ is solved, the phase for the native data set is achieved by a simple vector difference $\underline{F}_{PH}^{\lambda 1}$ - $\underline{F}_{H}^{\lambda 1}$ (Rhodes, 2006).

In order to use anomalous differences for MAD phasing, a fluorescence energy scan near the absorption edge of the anomalous scatterer is performed providing the proteins characteristic curves for f' and f'' (Figure 11). Typically three or four different wavelength data sets around the absorption edge are measured to get anomalous signals. In particular these wavelengths are: the maximum (peak) of f'', where the signal for the imaginary term shows its largest value, the minimum of f' corresponding to the inflection point of f'' and a wavelength far away from the absorption edge (low and/or high energy remote). In the end one obtains for each wavelength a set of different anomalous signals, which can be used for phase information and handedness. For this method, however, all data sets need to be recorded from a single crystal, as the difference between Friedel mates caused by the anomalous scatterer is less than 5% of the entire signal.

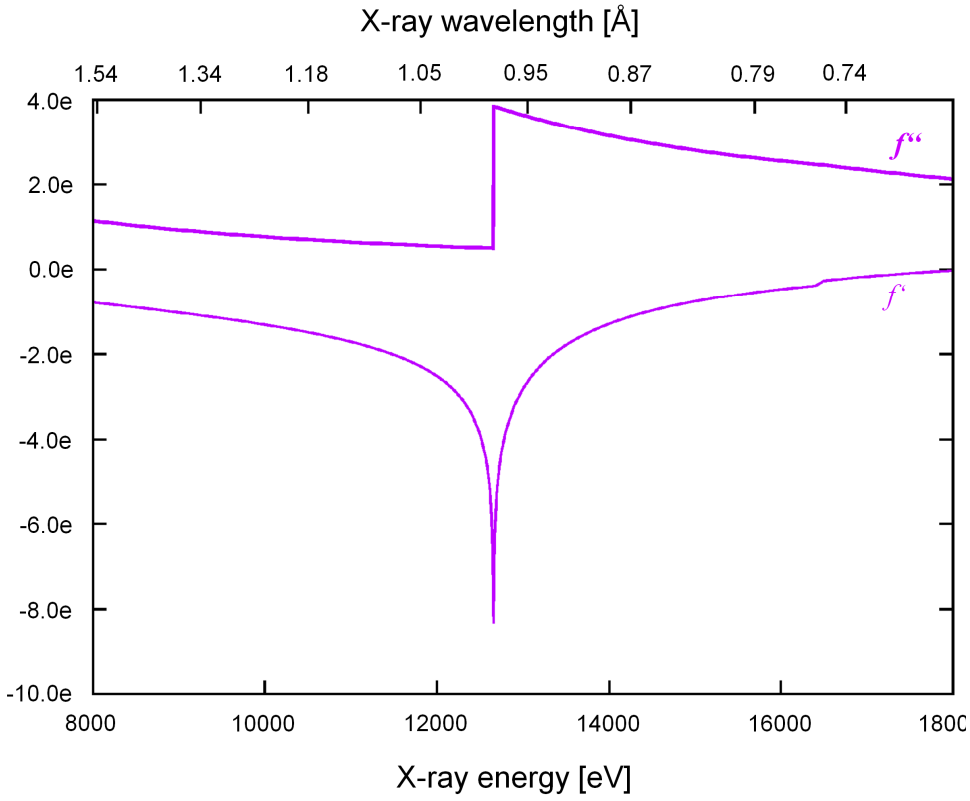


Figure 11: Theoretical plot of scattering factors f'' and f' of Se as a function of the X-ray energy and wavelength. Plot has been generated using the edge plot web tool (www.bmsc.washington.edu/scatter).

With the assistance of the measured anomalous differences or Bijvoet differences (difference between the measured amplitudes of the Friedel mates at one wavelength) and the dispersive differences (difference between the measured amplitudes of the Friedel mates at two different wavelengths) Patterson maps can be calculated for the anomalous scatterer. Both maps should contain the same peaks for solving the heavy atom substructure. From the initial positions of the anomalous scatterers the structure factor of $\underline{F}_{PH}(+)$ and $\underline{F}_{PH}(-)$ can be deduced graphically by a Harker construction (Figure 10). However, this method is only a graphical illustration, computer programs determine phase angles by the MAD observational equations (Hendrickson *et al.*, 1985). In the case of MAD experiment using two different wavelengths (λ_1 and λ_2) four equations can be set up (Equation 4). Accordingly six or eight equations can be set up, if three or four wavelengths are used. As the structure factor amplitudes of F_{PH} and F_P are wavelength independent, the three unknown parameters F_{PH}, F_P and (α_{PA}-α_P) are over-determined by the four, six or eight equations. If the heavy atom substructure is already known, the protein phase angles can be good estimated.

$$
\begin{aligned}
{}^{\lambda 1}F(+)^2 &= F_{PA}^2 + a(\lambda_1)F_A^2 + b(\lambda_1)F_{PA}F_A\cos(\alpha_{PA}-\alpha_A) + c(\lambda_1)F_{PA}F_A\sin(\alpha_{PA}-\alpha_A)\\
{}^{\lambda 1}F(-)^2 &= F_{PA}^2 + a(\lambda_1)F_A^2 + b(\lambda_1)F_{PA}F_A\cos(\alpha_{PA}-\alpha_A) - c(\lambda_1)F_{PA}F_A\sin(\alpha_{PA}-\alpha_A)\\
{}^{\lambda 2}F(+)^2 &= F_{PA}^2 + a(\lambda_2)F_A^2 + b(\lambda_2)F_{PA}F_A\cos(\alpha_{PA}-\alpha_A) + c(\lambda_2)F_{PA}F_A\sin(\alpha_{PA}-\alpha_A)\\
{}^{\lambda 2}F(-)^2 &= F_{PA}^2 + a(\lambda_2)F_A^2 + b(\lambda_2)F_{PA}F_A\cos(\alpha_{PA}-\alpha_A) - c(\lambda_2)F_{PA}F_A\sin(\alpha_{PA}-\alpha_A)
\end{aligned}
$$

Equation 4

$$a(\lambda_n) = (f'(\lambda_n)^2 + f''(\lambda_n)^2)/f_0^2$$

with $$b(\lambda_n) = 2f'(\lambda_n)^2/f_0$$

$$c(\lambda_n) = 2f''(\lambda_n)^2/f_0$$

3.7.2 Model building, refinement and validation

Once the phase problem has been solved successfully electron density has to be interpreted by building a protein model. The initial model building can be done automatically, if data quality and resolution are sufficiently good. The final model building is performed manually to improve the model with regard to the chemical and biological properties.

The principal aim of a protein crystal structure determination is the building of a reasonable model that is in good agreement with the observed data. For this purpose the positions of individual atoms are optimized in iterative rounds of refinement calculations, that improve the overall model and hence the model-derived phases. The fit in quality of the model to the observed data is measured by the crystallographic R factor R_{cryst}, which can be determined from the calculated model structure factors F_{calc} and the observed structure factors F_{obs} (Equation 5) . Well defined structures generate R_{cryst} values of 10-25 %, whereas randomly placed atoms in the unit cell have a maximum value of 59 % (Wilson, 1949). In order to prevent over-fitting, a test set of data (usually 5-10 % of all reflections) is not used for refinement calculation. Using the test set of reflections, the so-called R_{free} factor can be calculated (Brunger, 1993). If the R_{free} value increases, after changing the model parameters, a model bias has been generated. Finally the model quality is validated by stereo chemical parameters (Engh & Huber, 1991), such as bond length, bond angles and torsion angles (Ramachandran plot) (Ramachandran & Sasisekharan, 1968), the reliability of the R factors R_{cryst}/R_{free} and a contact analysis of all atoms (Lovell *et al*., 2003).

$$R_{cryst} = \frac{\sum_{hkl}\left|F_{obs}(hkl) - k\cdot F_{calc}(hkl)\right|}{\sum_{hkl}\left|F_{obs}(hkl)\right|}$$

Equation 5

with k the resolution-dependent scale factor

3.7.3 Structure determination of $YscU^C$

Four multiwavelength anomalous diffraction (MAD) data sets of SeMet-labelled $YscU^C$ N263A were collected at beamline X12 at the EMBL Outstation (Hamburg, Germany) and processed with the HKL2000 (Otwinowski, 1997) and MOSFLM (Leslie, 1992). Before collecting a whole data set, test pictures were taken to calculate the best data collection strategy. The solution of the selenium substructure, density modification and partial model building were done with the automated crystal structure determination platform: Auto-Rickshaw (Panjikar *et al*., 2005) at the EMBL-Hamburg. Before starting the routine, the following input data was prepared. The four MAD data sets were processed anomalously with HKL2000 (Otwinowski, 1997). The obtained intensities were converted into structure factors in a MTZ file using the program SCALEPACK2MTZ (French & Wilson, 1978). The created MTZ files and the amino acid sequence of $YscU^C$ were applied to the interface and a four wave length MAD experiment was chosen. The program estimated from the files the unit cell dimensions, the number of molecules per asymmetric unit (1) and prepared an input file for SHELX programs. Once the program SHELX C (Pape & Schneider, 2004) had solved the MAD equations for the anomalous scatterer, an input file was written up to give a useful resolution range. The subsequent program SHELX D (Schneider & Sheldrick, 2002) solved the selenium substructure. Finally SHELX E (Pape & Schneider, 2004) solved the phasing equations. The correct substructure enantiomorph ($P4_32_12$) was estimated by several cycles of density modifications. Here the contrast of the electron density map (heavy-atom) of the correct enantiomorph is much higher. After localization of two selenium sites a partial model was generated with ARP/wARP (Morris *et al.*, 2003). Finally the model was completed manually using COOT (Emsley & Cowtan, 2004) and refinement was carried out using REFMAC5 (Murshudov *et al*., 1997).

The X-ray data set of $YscU^C$ N263D was collected at ESRF beamline ID29 (Grenoble, France), processed with MOSFLM (Leslie, 1992) and scaled with SCALA (Evans, 2006). Using REFMAC5 (Murshudov *et al*., 1997) and SeMet-labelled $YscU^C$ N263A as a phasing model, the structure was solved by difference Fourier and further refined with COOT (Emsley & Cowtan, 2004) and REFMAC5 (Murshudov *et al*., 1997).

The validation of both structures was performed with MOLPROBITY (Davis *et al*., 2007). Data collection and refinement statistics are given in Table 7. The coordinates have been deposited in the Protein Data Bank (PDB) under accession codes 2v5g and 2w0r.

3.8 Protein nuclear magnetic resonance spectroscopy

In addition to X-ray crystallography protein nuclear magnetic resonance spectroscopy (NMR) is a further important method for elucidating the 3D structures of small proteins. A great advantage of protein NMR is the possibility of analyzing protein samples in solution. Further, once the protein structure is solved, interactions can be investigated more easily. A drawback is the size limitation of investigating proteins (~ 30 kDa) due to overlapping signals.

Nuclei (e.g ^{1}H, ^{19}F or ^{31}P), which contain a nuclear spin, exist in discrete nuclear spin states when they reside in an external magnetic field. NMR spectroscopy uses radiofrequency radiation to observe transitions between the different nuclear spin states. Consequently, nuclei that do not have a nuclear spin are transparent to NMR spectroscopy. However, this concerns the most common isotopes like ^{12}C (98.9 %), ^{14}N (99.6 %) and ^{16}O (99.8 %) nuclei. Due to the high natural abundance of these common isotopes, protein samples have to be labelled with isotopes possessing a nuclear spin. For this purpose proteins can be produced recombinantly in minimal media containing labelled ^{13}C-labelled glucose or ^{15}N-labelled ammonium chloride.

If all nuclei absorbed energy at the same resonance frequency, NMR spectroscopy would not be very useful for investigations. However, as electrons of the molecule create a local weak magnetic field, the nuclei are shielded from the applied main field. Therefore the nuclei absorb, depending on their local environments, at slightly different frequencies, resulting in a so-called chemical shift in the NMR spectrum. The interpretation of a 1D spectrum is not possible for complex molecules such as proteins, as too many signals are overlapping. Therefore the spectrum is expanded to a second or third dimension by correlation of the resonance frequencies to the frequencies of same or hetero nuclei. The resulting peaks are mostly separated allowing an assignment and novel information about the chemical environment (Lottspeich & Engels, 2006).

3.8.1 Investigation of conformational change with protein NMR

Comparing two different protein states (e.g. before and after cleavage) can be easily performed by recording a HMQC (heteronuclear multiple quantum coherence) spectrum of both states (Gal *et al*., 2007). In this 2D spectrum nitrogen frequencies are correlated with the amide proton frequencies. Every observed peak represents a proton bound to its ^{15}N-atom. This means the protein spectrum consists basically of peaks from H^{N}-protons of backbone and side chain amids. A conformational change in the protein changes the chemical environment of the amide protons thus leading to different chemical shifts. From superposition of both spectra it is clear whether there is a conformational change or not.

3.8.2 The ^{15}N, ^{1}H HMQC experiment

Protein NMR measurements were performed by Dr. Thorsten Lührs (HZI, Braunschweig). ^{15}N-labelled $YscU^{C}$ variants were expressed and purified as described in Chapter 3.3.1. For protein NMR spectroscopy, we used a 12 mg/mL solution of uniformly ^{15}N-labelled $YscU^{C}$ in a mixed solvent of 95% (v/v) H_2O, 5% (v/v) $^{2}H_2O$ and a 10 mg/mL solution of $YscU^{C}$ N263A in the same solvent. The NMR samples contained 50 mM NaCl, 3 mM KCl, 12 mM Na_2HPO_4 and 2 mM KH_2PO_4 at pH 6.8. The Bruker Advance III 600 MHz NMR spectrometer used for this study was equipped with a 5-mm Z-axis gradient triple-resonance cryo-probehead (providing a better signal noise ratio). The 2D [^{15}N,^{1}H] correlation spectra were recorded at ω_1 = 2100 Hz and ω_2 = 8400 Hz. The maximal evolution times were t_{1max} = 60 ms and t_{2max} = 240 ms and the time domain data size was 256 × 2048. Programs PROSA (Guntert *et al*., 1992) and CARA (Keller, 2004) were used for data processing and spectral analysis, respectively.

3.9 Bioinformatics and figure representation

Plasmids, genes and proteins were analyzed and archived in Vector NTI version 8. Pair wise and multiple sequence alignments were performed using CLUSTALW (Thompson, 1994). Figures for protein representations were prepared using PYMOL (http://pymol.org).

3.10 Molecular orbital diagram

A molecular orbital diagram (MO diagram) is a qualitative descriptive tool explaining chemical bonding in molecules. Accordingly atomic orbitals (AO) and/or molecular orbitals (MO) are combined using the linear combination of atomic orbitals (molecular orbital) method (LCAO method). The simplest case displays the σ-bound of two H-atoms (Figure 12). Combination of both AO yield an occupied binding σ-MO (lower energy) and an unoccupied anti-binding σ*-MO. In a corresponding more complex "hybrid MO" the corresponding MOs can also be combined (Riedel & Janiak, 2002).

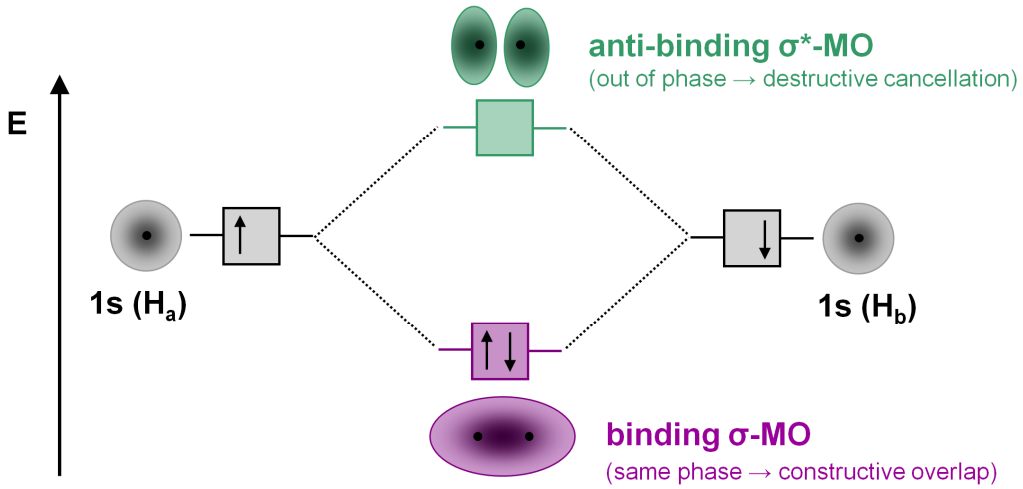


Figure 12: Energy diagram of the H_2 molecule. LCAO of two 1s-orbitals of H-atoms results in a binding and anti-binding MO. At the ground state both electrons pair in the binding σ-MO according to a σ-bound.

4 Results

4.1 The type III secretion recognition protein YscU

4.1.1 Design of the YscU^C construct

Initially, a suitable cytoplasmic YscU construct and the corresponding expression vector had to be designed for recombinant expression of YscU in *E.coli* cells. As YscU from *Yersinia enterocolitca* (Swiss-Prot. accession number Q56844) consists of an N-terminal trans-membrane and a C-terminal cytoplasmic domain (Sorg *et al*., 2007), a prediction of the domain borders was made by using the topology prediction server PHOBIUS (Kall *et al*., 2004).

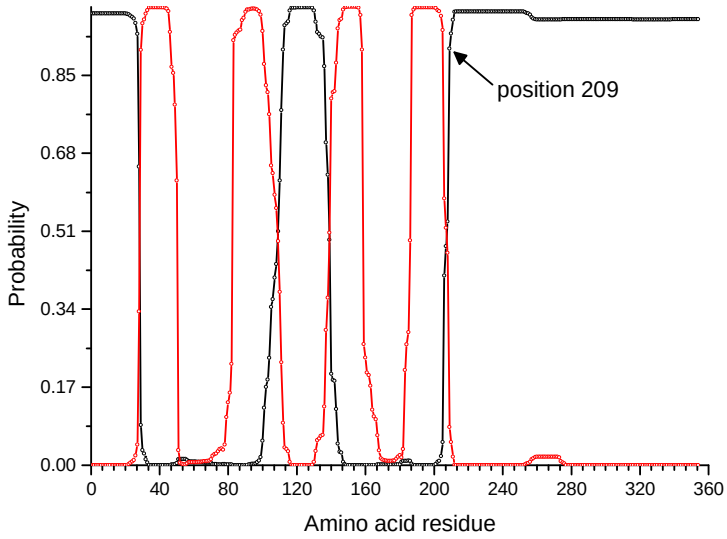


Figure 13: Topology and trans-membrane plot of YscU. The plot shows the probabilities of cytoplasmic (black) and TM helix (red) regions of YscU amino acid residues.

The plot (Figure 13) shows four transmembrane helices in the N-terminal domain. The domain ends roughly at amino acid 208. Amino acids 211-354 were chosen to obtain a construct containing just YscU^C. The corresponding *yscU^C* gene was cloned into a pGEX-6P-1 vector using *BamHI* and *AvaI* restriction sites. Gene expression yields a GST-YscU^C fusion protein, which is cleavable by the PreScission protease. For further experiments several variants of YscU^C were generated by using site-directed mutagenesis.

4.1.2 Preparation of YscU^C

For wild type YscU^C and all YscU^C variants a protein preparation scheme (Figure 14) was established to obtain sufficient amounts of pure protein. The protein preparation scheme is shown exemplarily for the wild type protein.

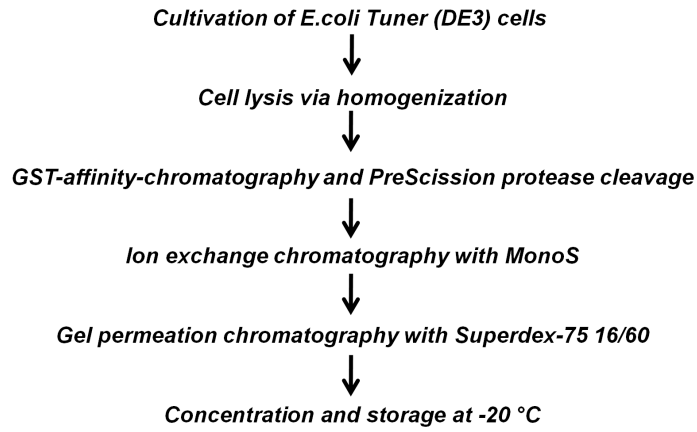


Figure 14: Protein preparation scheme for YscU^C and its variants.

After cultivation and cell lysis of GST-YscU^C producing *E. coli* cells (see Chapter 3.3.1), the soluble supernatant was applied to glutathione sepharose for initial purification (Figure 15). GST-YscU^C bound to the sepharose matrix, whereas unbound protein was removed by three washing steps. The GST fusion tag was cleaved off by adding PreScission protease and remained on the sepharose matrix, while YscU^C had been eluted. The SDS-PAGE shows YscU^C as a double band (10.5 and 6.7 kDa) due to its autocleavage activity. Subsequently, YscU^C containing fractions were dialyzed against IEC-buffer A to remove salt excess.

The second purification step involved a cation exchange chromatography (MonoS 10/100 column). Therefore, runs were optimized to two linear step gradients: 1st 0-30 % IEC buffer B in 3 column volumes and 2nd 30-100% IEC buffer B in 1 column volume. After all protein solution had been applied to the column, YscU^C was eluted (Figure 16) as a concentrated single peak in the first gradient at ~19 % IEC buffer B (~ 20 mS/cm).

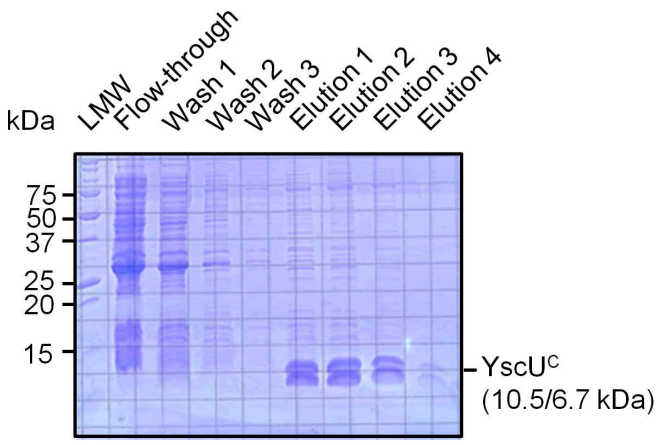


Figure 15: SDS-gel of the GST affinity chromatography of YscU^C. Autocleavage cleave YscU^C into a 6.7 kDa N-terminal fragment and a 10.5 kDa C-terminal fragment.

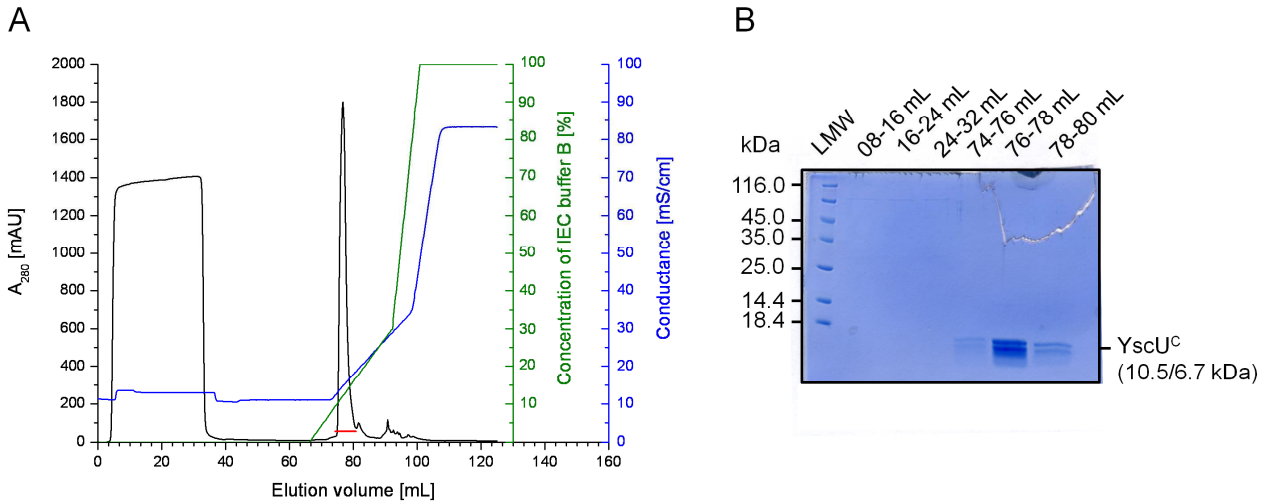


Figure 16: IEC elution profile (A) and corresponding SDS-gel (B) for YscU^C. Pooled YscU^C fractions are depicted as a red bar in the elution profile.

As a final purification step gel permeation chromatography was performed using a Superdex 75 16/60 prep grade column (GE Healthcare). The elution profile (Figure 17) shows a single peak at 71.5 mL. However, the peak is not fully symmetrical and shows a small tailing preceding the main peak. Sometimes the tailing also results in a

small additional peak containing YscU^C. The YscU^C elution volume corresponds to a mass of approximately 28 kDa compaired to a calibration curve (Figure 18). A subsequently performed DLS experiment showed a mono-disperse sample with a hydrodynamic radius of 2.36 nm (15.7 % polydispersity) and a calculated mass of 25.1 kDa for YscU^C. As both methods are based on the assumption of a globular shaped protein, the obtained masses are generally too high and therefore a monomer for YscU^C is assumed. After validation of each fraction by SDS-PAGE, pure peak fractions were pooled and concentrated. For future experiments the protein was stored at -20 ℃ and aliquoted in vials containing 100 µL protein solution in GPC buffer at a concentration of 10-20 mg/mL.

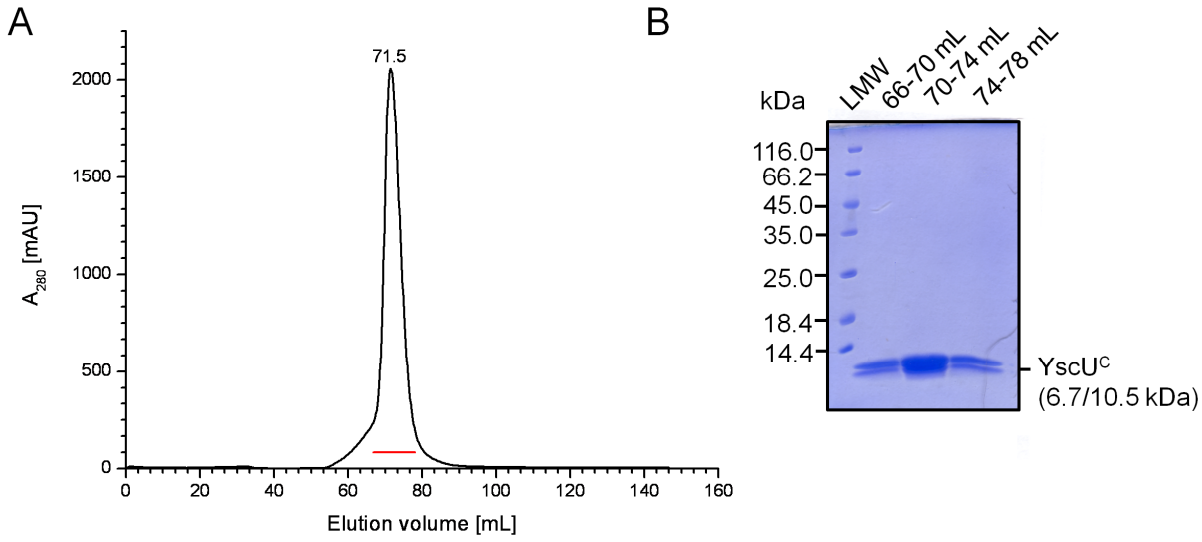


Figure 17: GPC elution profile (A) and corresponding SDS-gel (B) for YscU^C. Pooled YscU^C fractions are depicted as a red bar in the elution profile. The calculated mass of YscU^C fram a GPC run is approximately 28 kDa.

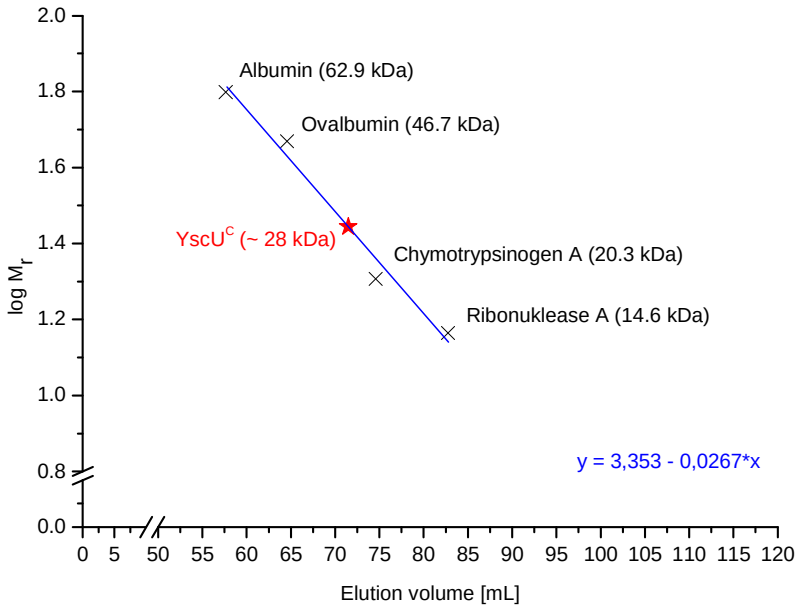


Figure 18: Calibration curve of Superdex 75 16/60 and linear fit. The calibration run was performed using four different calibration proteins and a flow rate of 0.8 mL/min. Using the linear equation the molecular weight of YscU^{C} (71.5 mL) corresponds to approximately 28 kDa.

In summary, wild type YscU^{C} and its variants were obtained in high purity and yield about 10-15 mg per L cell culture. However, SeMet – and ^{15}N-labelled YscU^{C} variants yielded only 3-5 mg/mL, due to the use of minimal medium. Successful SeMet incorporation was proven by mass spectroscopy (Figure 19). Each variant was checked for degradation, aggregation and freezing problems using DLS and SDS-PAGE. All variants turned out to be stable even after long term storage.

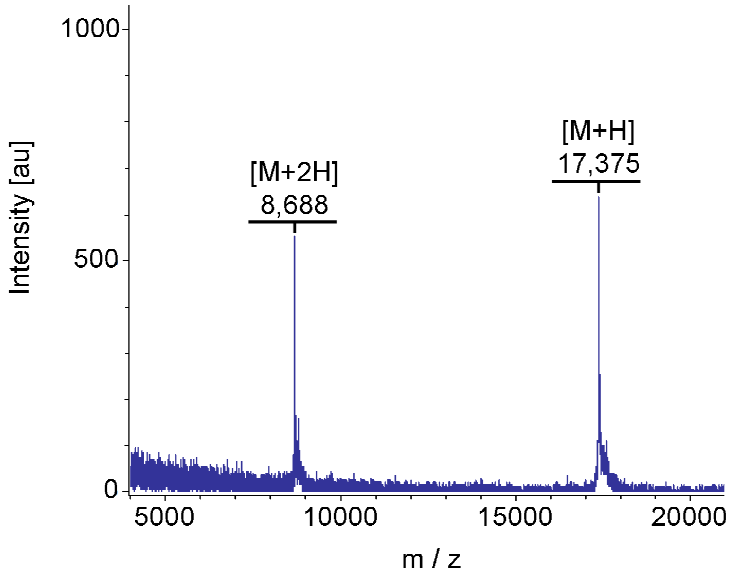


Figure 19: MALDI-TOF-MS spectrum of purified SeMet YscU^C N263A variant. [M+H] indicates the singly charged species, [M+2H] the doubly charged ion. The calculated molecular mass of fully SeMet-labeled YscU^C N263A is 17,378 kDa. The molecular mass also includes the N-terminal amino acids GPLGS that are remained on the protein after PreScission cleavage.

4.1.3 Crystallization of YscU^C

Despite intensive screening using a variety of different protein concentrations, temperatures and additives, no crystals could be obtained for wild type YscU^C. In contrast the YscU^C N263A variant readily crystallized at 20 ℃ from a crystallization solution containing 1.6 M $(NH_4)_2SO_4$, 0.2 M NaCl and 0.1 M Hepes (pH 7.5). The octahedral shaped crystals (Figure 20) grew within 4 days. Optimization by varying the pH and the $(NH_4)_2SO_4$ concentrations yielded in larger crystals with sizes up to 450 µm x 450 µm x 800 µm. Under the same crystallization conditions crystals with the same habitus were obtained from the variants YscU^C N263D and SeMet labelled YscU^C N263A (Figure 20). All crystals were transferred into cryo-buffer containing reservoir buffer with 16 % (v/v) glycerol as cryo-protectant and flash-frozen at 100 K in liquid nitrogen.

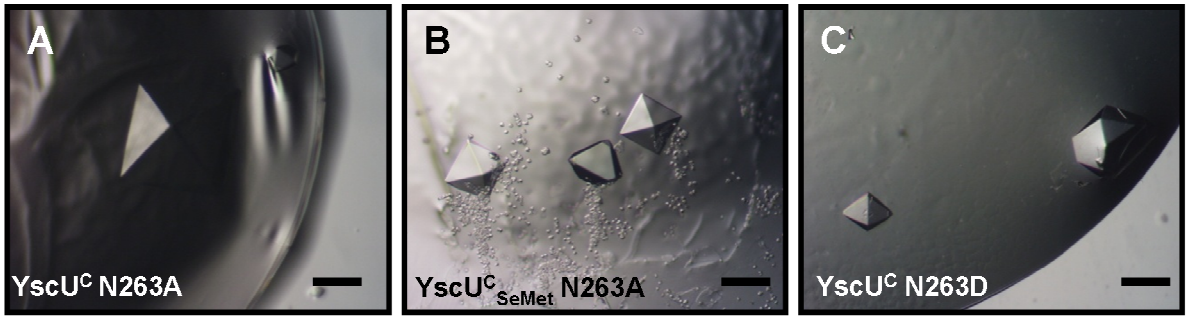


Figure 20: Crystals of different YscU^{C} variants. The scale bar corresponds to 200 µm.

4.1.4 Data collection and space group determination

Data collection for X-ray data sets of YscU^{C} N263D and SeMet labelled YscU^{C} N263A were performed at the ESRF beamline ID29 (Grenoble, France) and at the wavelength tuneable beamline X12 at the EMBL Outstation (Hamburg, Germany), respectively. For all isomorphous crystals a tetragonal symmetry with unit cell dimensions of a = b = 66.2 Å and c = 68.1 Å was determined. Analysis of systematic absence revealed missing reflections for the following conditions: along the crystallographic *l* axis with *l*=4n (00*l*, n=integer) and the *h* axis with *h*=2n (*h*00, n=integer). This indicated two possible enantiomorphous space groups: $P4_12_12$ or $P4_32_12$. Matthews probability analysis (Kantardjieff & Rupp, 2003; Matthews, 1968) revealed a packing parameter V_M = 2.17 Å³/Da and a solvent content of 45 %. As only one molecule per asymmetric unit was estimated for unit cell content, an analysis for non crystallographic symmetries was not necessary. However, a self rotation function (Tong & Rossmann, 1997) was calculated (data not shown) and the data confirmed the Laue crystal class of 4/mmm. Data collection statistcs are summarized in Table 7 at the end of the Chapter 4.1.6.

4.1.5 Fluorescence scan for MAD phasing

As there was no appropriate protein model available for Molecular Replacement, phasing was achieved using anomalous dispersion at multiple wavelengths (MAD-phasing). For this reason SeMet labelled YscU^{C} N263A crystals were analyzed using

a fluorescence scan (Figure 21). The fluorescence spectrum near the selenium K-edge revealed a protein specific absorption curve as a function of energy. The specific absorption edge provides the necessary information for finding f'' and f' and in turn the appropriate wavelength for measuring anomalous diffraction data. A standard fluorescence scan was performed between 12668 eV and 12728 eV and the following wavelengths were chosen for anomalous data set collection: 0.97853 Å (peak), 0.97892 Å (inflection point), 0.95369 Å (high energy remote) and 1.0000 Å (low energy remote). f'' and f' were calculated (Figure 22) from the fluorescence scan using the program CHOOCH (Evans & Pettifer, 2001).

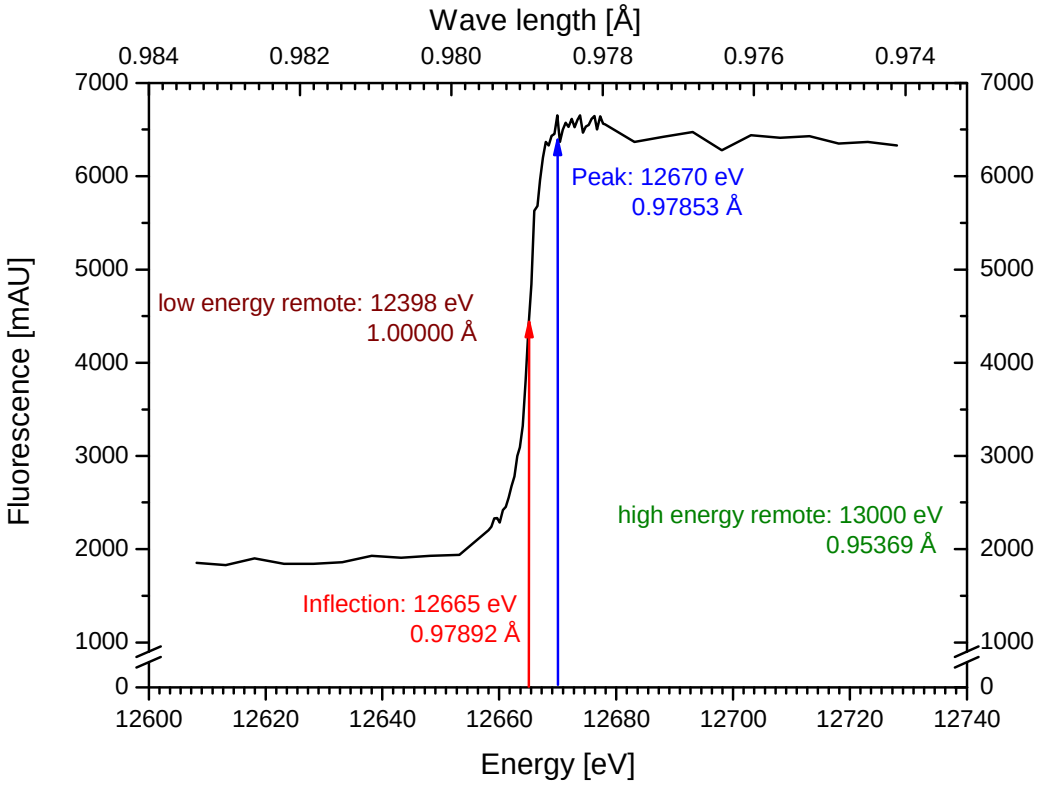


Figure 21: X-ray fluorescence scan for SeMet labelled YscU^{C} N263A. The scan was performed at the K absorption edge of selenium. The chosen wavelengths for MAD data collection are shown in the figure.

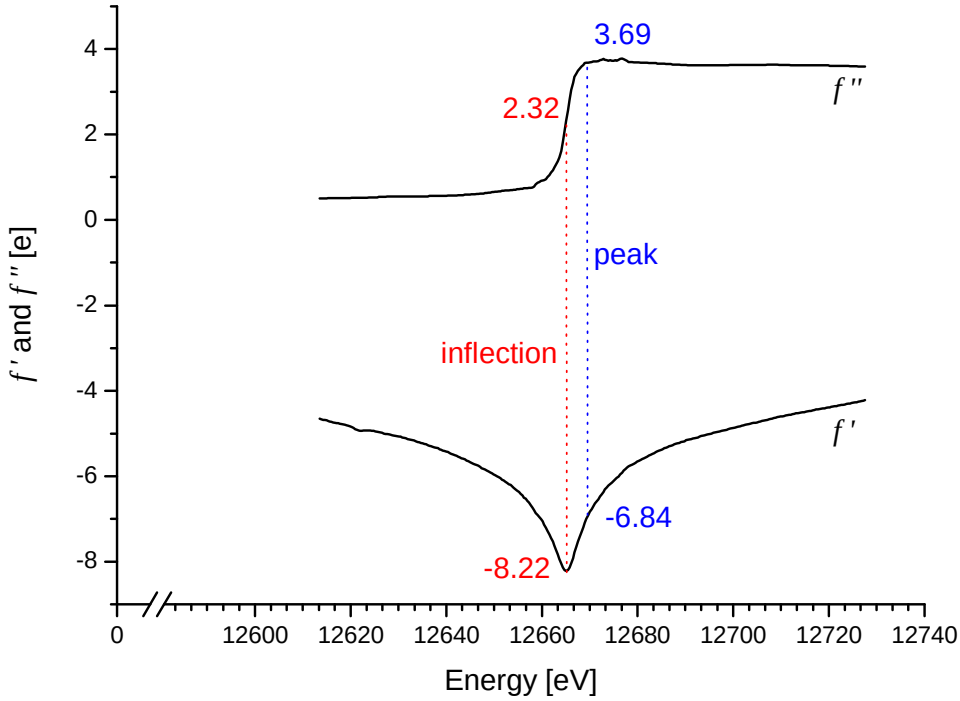


Figure 22: Plot of the YscU^C N263A specific scattering factors *f''* and *f'* as a function of the X-ray energy. Specific values of *f''* and *f'* at the inflection point and the peak are depicted in red and blue, respectively. The curves were calculated from the fluorescence scan using the program CHOOCH (Evans & Pettifer, 2001).

4.1.6 Structure determination of YscU^C

Structure determination of YscU^C N263A

The solution of the selenium substructure, density modification and partial model building were done with the automated crystal structure determination platform: Auto-Rickshaw (Panjikar *et al.*, 2005) at the EMBL-Hamburg outstation. The system combines several macromolecular computer programs from the CCP4 suite (Collaborative Computational Project, Number 4, 1994) into an automated software pipeline solving macromolecular structures. Therefore, the platform allows for rapid analysis of structure solution and validation already at the synchrotron beamline. However, the process has to be carefully checked at different stages based on the

obtained output. The obtained initial electron density map revealed the positions of the selenium sites (Figure 23). The two, out of the four expected selenium, sites were sufficient to solve the structure. The resulting electron density was of sufficient quality to allow construction of an initial model using the program Arp/wARP (Morris *et al*., 2003). In iterative cycles model building was carried out manually using COOT (Emsley & Cowtan, 2004) and refinement was completed using REFMAC5 (Murshudov *et al*., 1997).

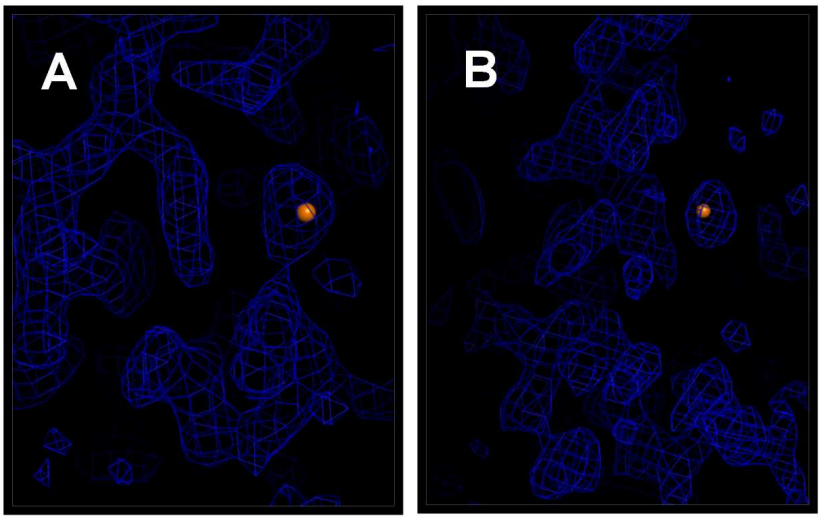


Figure 23: Initial electron density map of SeMet YscU^{C} N263A after density modification. The map is drawn at a contour level of 1.5 σ. The two detected selenium atom are depicted as orange spheres (A and B).

Structure determination of YscU^{C} N263D

Using the program PHASER (McCoy *et al*., 2005) and the refined structure of SeMet labelled YscU^{C} N263A as a phasing model, the structure of YscU^{C} N263D was solved using the difference Fourier technique. Final model building was performed manually using COOT (Emsley & Cowtan, 2004) and refinement was completed using REFMAC5 (Murshudov *et al*., 1997).

Table 7: Data collection and refinement statistics.

Data sets	*SeMet-labeled YscU^C N263A*				*YscU^C N263D*
	Peak	***Inflection***	***High-energy remote***	***Low-energy remote***	
Data collection ststistics					
Beamline	DESY X12	DESY X12	DESY X12	DESY X12	ESRF ID29
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit cell dimensions					
a (Å)	66.2	66.2	66.2	66.3	66.5
c (Å)	68.1	68.1	68.1	68.8	68.1
Wavelength (Å)	0.97854	0.97893	0.95370	1.0000	0.9500
Resolution (Å)	30.0-2.33 (2.41-2.33)	40.0-2.33 (2.41-2.33)	40.0-2.33 (2.41-2.33)	30.0-2.0 (2.11-2.0)	10.0-1.55 (1.63-1.55)
Mosaicity (°)	0.80	0.98	0.89	0.95	0.24
Unique reflections	6910 (670)	6907 (663)	6920 (667)	10836 (1517)	21871 (3232)
Completeness (%)	100 (99.9)	100 (100)	100 (100)	100 (100)	96.7 (99.9)
Multiplicity	13.6 (12.0)	13.6 (12.6)	13.8 (13.3)	13.6 (13.6)	3.1 (3.1)
I/σI	31.7 (8.2)	34.3 (8.8)	34.0 (9.5)	8.1 (2.2)	8.4 (4.8)
R_{merge}(%)	8.2 (31.5)	7.7 (29.8)	7.8 (31.0)	6.9 (35.2)	5.2 (15.1)
Wilson B-factor (Å^2)	39	39	38	25	19
Solvent content (%)	45	45	45	45	45
Refinement statistics					
R_{cryst} (%)				22.1	20.0
R_{free} (%)				25.2	22.2
Number of atoms					
Protein				1021	1104
Solvent				56	135
Cl^-				1	1
r.m.s.d. from ideal					
Bond lengths (Å) /				0.015	0.013
Bond angles (°)				1.701	1.502
Ramachandran plot regions					
Favoured (%)				95.3	96.4
Allowed (%)				4.7	3.6
Average *B*-factor					
Protein (Å²)				29.7	20.2
Solvent (Å²)				35.5	33.9
Cl^- (Å²)				27.0	21.9

Values in parentheses refer to shell of highest resolution. R_{free} test set size for YscU N263A 5% and for YscU N263D 3%.

4.1.7 Validation of the YscU^C structures

The final refinement statistic is given in Table 7. The quality of both protein models was checked regularly during refinement on the basis of the R_{cryst}/R_{free}-values and geometric parameters during refinement. Both structures were refined to R_{cryst}/R_{free}-values of 22.1/25.3 % for SeMet YscU^C N263A and 20.1/22.5 % for YscU^C N263D, respectively. The protomers, present in the asymmetric unit cell together with solvent molecules, were well defined for the amino acid residues 222-342 in SeMet YscU^C N263A and for amino acid residues 222-343 in YscU^C N263D. The remaining N- and C-terminal amino acids were not traceable in the electron density map and probably flexible.

Bond lengths and bond angles showed no significant deviations from standard structures (Engh & Huber, 1991). Validation of the torsion angles and an all-atom contact analysis were carried out using the programs RAMPAGE (Lovell *et al*., 2003) and MOLPROBITY (Davis *et al*., 2007). In both structure all residues grouped in the most favoured or allowed regions of the Ramachandran plot (Figure 24).

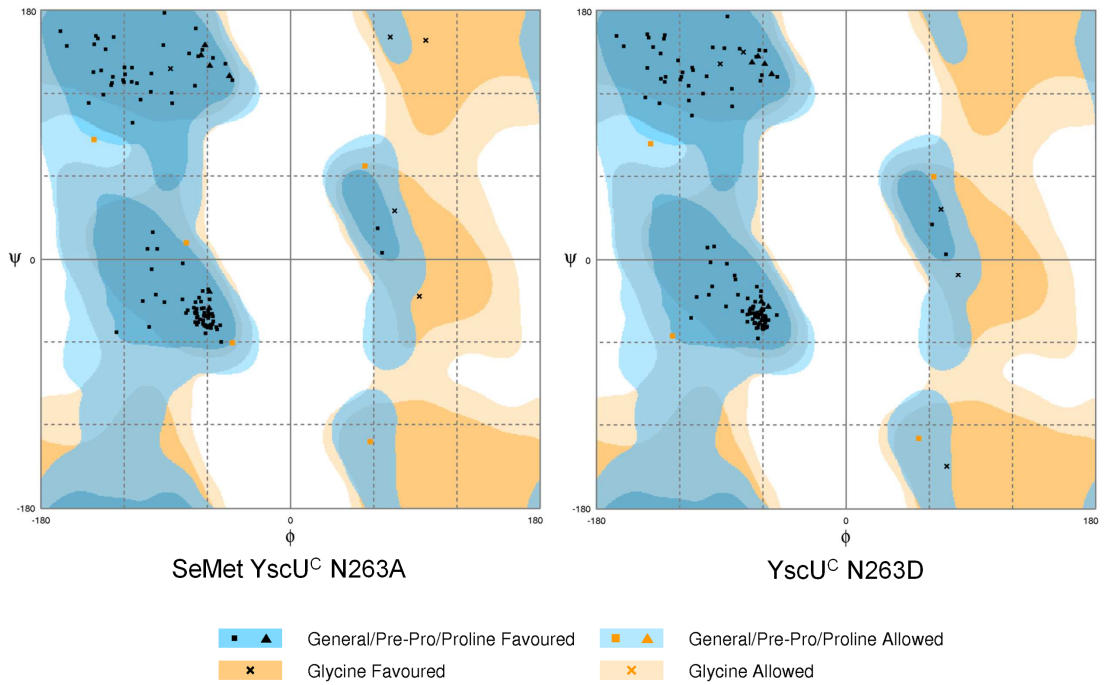


Figure 24: Ramachandran plot of SeMet YscU^C N263A and YscU^C N263D. Favoured and allowed regions are coloured dark and light, respectively. Forbidden regions are shown in white. The plot was obtained using the RAMPAGE (Lovell *et al*., 2003) server.

Good agreement of protein model and its electron density was indicated by the real-space correlation coefficient (Rupp, 2006), which was calculated using the program OVERLAPMAP (CCP4 suite, 1994). In Figure 25 the main chain coefficients were well defined for SeMet YscU^{C} N263A, whereas the coefficients for YscU^{C} N263D are generally weaker, especially at the N- and C-terminus, the loop connecting β strands 2 and 3 and the loop connecting α-helix 2 and β strand 4. These weak coefficients are most likely a consequence of the rather harsh geometric restraints applied during refinement at high resolution. The *B*-factor analysis is also given in Figure 25 and agrees well with the correlation coefficient. As expected, higher *B*-factors are observed at the loop regions and at the termini.

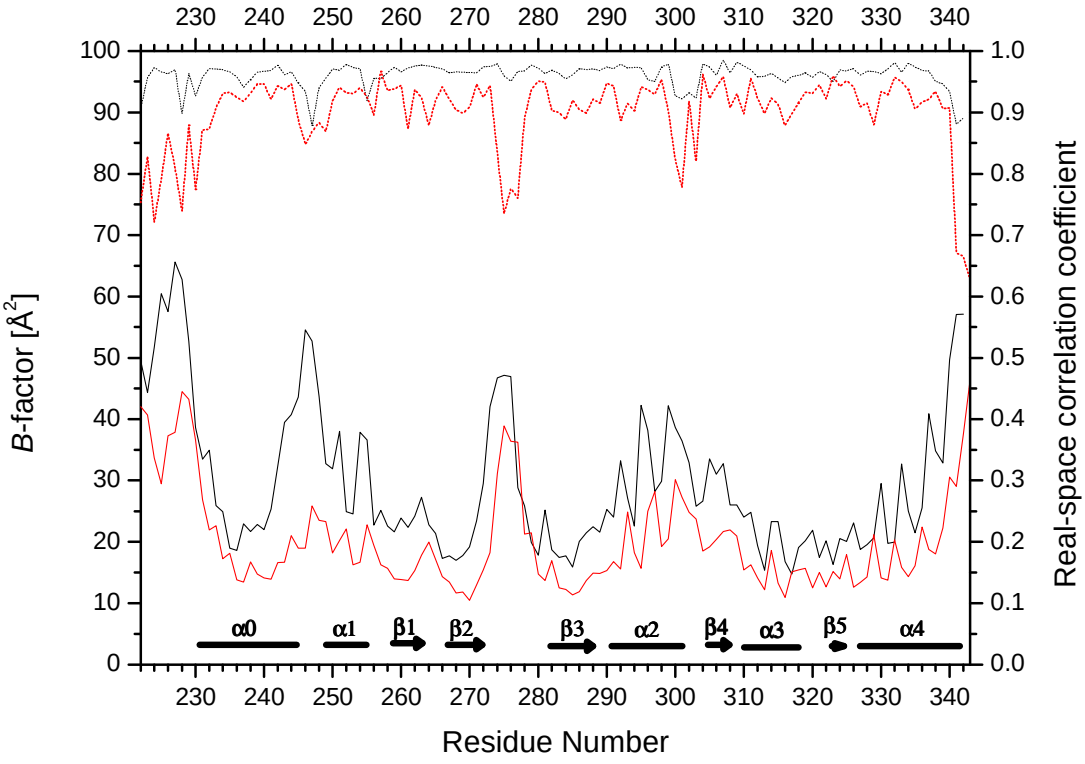


Figure 25: Plot of main chain residuals of B-factors and real-space coefficient correlation. SeMet YscU^{C} N263A and YscU^{C} N263D are depicted in black and red, respectively. B-factor values are drawn as a full line and the real-space correlation coefficient is drawn as a dashed line. Secondary structure elements of SeMet YscU^{C} N263A are drawn at the bottom.

4.1.8 Crystal structure of uncleaved YscU^C

The globular domain of uncleaved YscU^C (Figure 26) consists of a central, five-stranded, mixed β-sheet (I-V), surrounded by four α-helices (1-4). α-helix 3 is almost perpendicular to the strand orientation of the central β-sheet. The NPTH-cleavage motif forms a type II β-turn connecting β-strands I and II with position 263 located at the C-terminal end of β-strand I.

The globular autocleavage domain is preceded by an extended N-terminal loop containing a well defined α-helix (α0) comprising residues 232-244. This helix forms an antiparallel interaction with helix 4 of a symmetry-related YscU^C monomer in the crystal (see next Chapter).

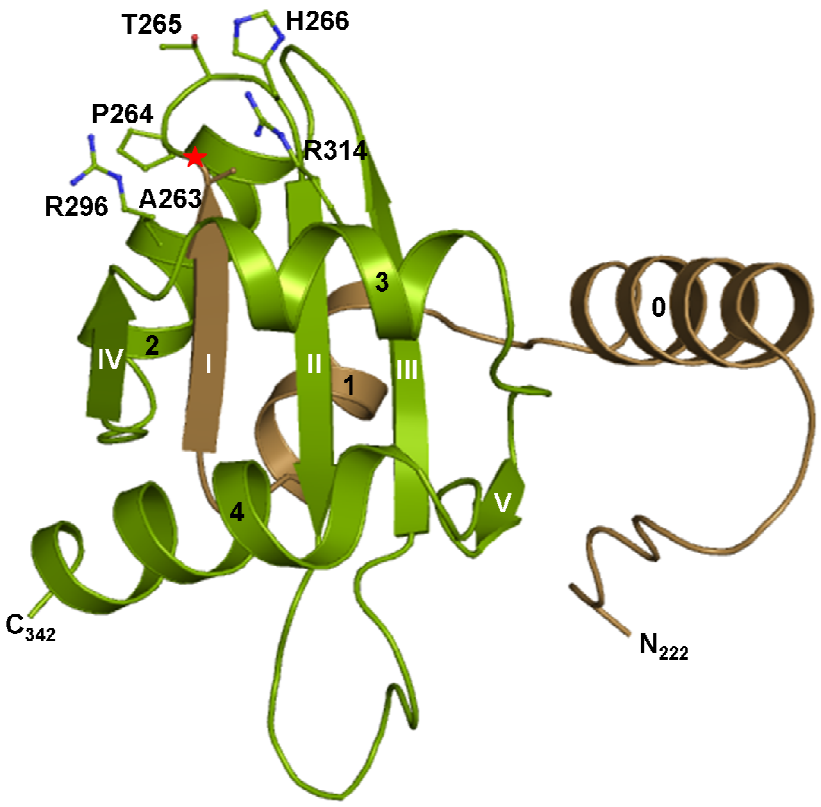


Figure 26: **Ribbon plot of the monomer of YscU^C N263A.** The position of the mutation in the cleavage site is indicated by a red star (*). After cleavage YscU^C is divided into an N-terminal and a C-terminal half, shown in brown and green, respectively. The modified NPTH motif, as well as Arg314 and Arg296, are depicted as ball-and-stick models.

4.1.9 Linker α-helix 0 is pinched due to crystal contacts

The crystal structure packing for uncleaved YscU^C was analysed using the PISA server (Protein interfaces, surfaces and assemblies service, (Krissinel & Henrick, 2007)). The result shows two large and tight crystal interfaces I and II (823 and 719 Å^2) and one small (227 Å^2) crystal interface III (Table 8). A detailed analysis of the interfaces (Figure 27) reveals strong interactions between α-helix 0 and α-helix 4 in interface II and between the NPTH motif plus a loop, connecting α-helix 2 and β-strand 3, in interface I. Obviously crystal contacts restricted the conformational freedom of α-helix 0 resulting in a well ordered and defined N-terminal part. In structure homologues of YscU^C the linker region is probably flexible and thus the linker domain is traceable in the electron density. The NPTH loop is involved in the crystal interfaces I and III, explaining why crystallization of cleaved, wild type YscU^C was not possible in this space group.

Table 8: Crystal contacts of YscU^C N263D.

Name	***Symmetry operation***	***Surface area [Å^2]***	***Interacting residues***	
			Molecule 1	***Molecule 2***
I	-y,-x,-z+1/2	823	232-233, 236-237, 239-240, 264-267, 286, 288-291, 293, 314, 317-319	232-233, 236-237, 239-240, 264-267, 286, 288-291, 293, 314, 317-319
II	-y+1/2,x-1/2,z-1/4	719	224-228, 273, 275-276, 278-281, 325-326, 328, 331-332, 334-335, 338-339, 342	222-223, 226, 229-231, 233-234, 237-238, 240-242, 244-247
III	x-1/2,-y-1/2,-z+1/4	227	265, 290, 293-294, 296-297	227-228, 328-329, 331

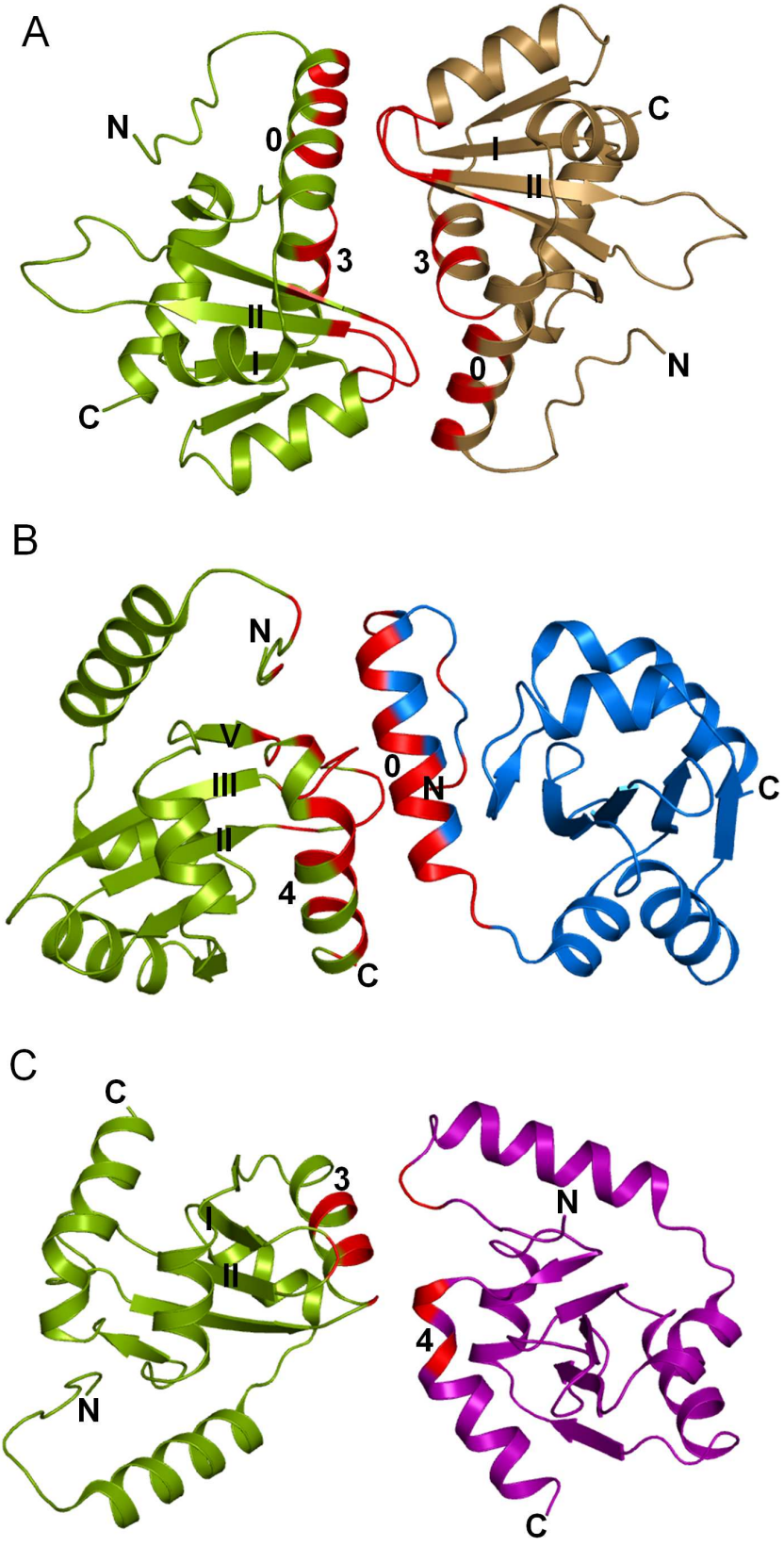


Figure 27: Ribbon plot of the crystallization interfaces between YscU^{C} monomers. Residues involved in interface interaction are drawn in red. The symmetry mate building (A) interface I is brown, - (B) interface II is blue and - (C) interface III is purple.

4.1.10 Structural basis for autocleavage of YscU

Close inspection of the modified NPTH motif in isosteric variant $YscU^C$ N263D (DPTH, Figure 28 A) shows that the main-chain carbonyl of Asp263 is tightly locked into position *via* hydrogen bonds to the backbone nitrogen atoms of His266 and Ile267. This structural stiffness is enhanced by the peptide bond with Pro264. All main-chain dihedral angles of the DPTH residues fall in the allowed regions of the Ramachandran plot, indicating no obvious conformational strain of main-chain atoms. A salt bridge with the guanidinium group of Arg314 positions the side chain of Asp263 above its main-chain carbonyl group.

The ideal entry angle for a nucleophilic attack of a carbonyl group has been proposed to lie between 100° and 110°, and a distance between nucleophile and electrophile of 2.5 Å. (Bürgi, 1974; Bürgi, 1973) The corresponding angle of the $YscU^C$ N263D variant is 109° with an N-C distance of 2.8 Å. The β-amide nitrogen of Asn263 would therefore be ideally positioned for a nucleophilic attack. Furthermore this geometric arrangement favours a stabilizing orbital overlap (Figure 28 B), where the anti-binding σ*-molecular orbital (MO; see Chapter 3.10) of the $C_{\alpha Asn}$-N_{Asn} bond interferes with the nascent binding σ-MO of the C(=O)-N-bond in native YscU. This entails a lowering of the activation energy that has to be applied for a new bond formation.

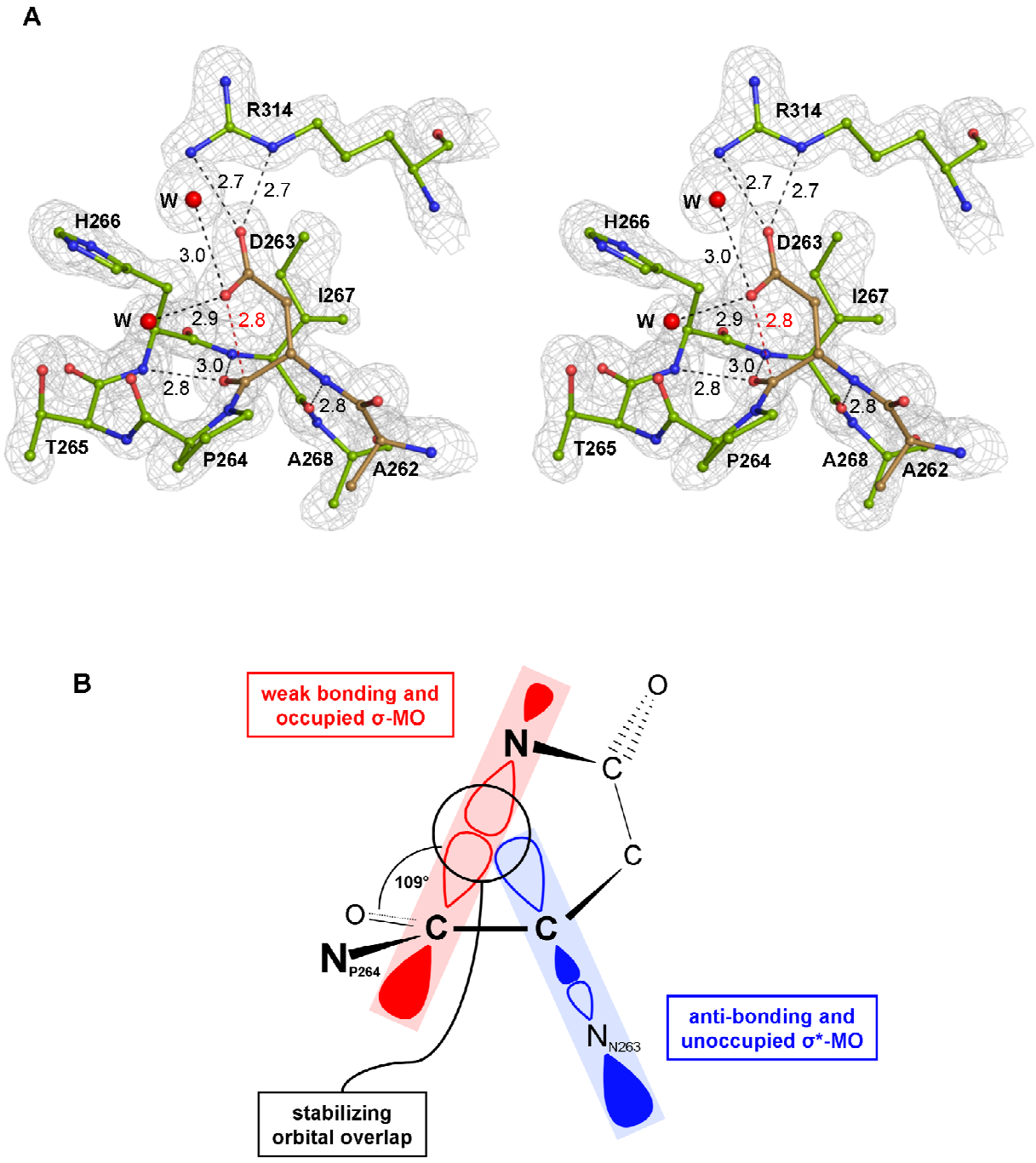


Figure 28: (A) Stereoview of the modified NPTH motif of YscU^{C} N263D. The C-terminal and N-terminal halves are shown in brown and green respectively. The isosteric Asp263 is optimally positioned (109°) for a nucleophilic attack, as indicated by the red dashed line. Hydrogen bonds are indicated by black dashed lines. The ($2F_O$-$F_{C,\alpha C}$) electron density map of the cleavage site is depicted at a contour level of 1.5 σ.

(B) Orbital-model illustrating the stabilizing orbital overlap (circle). The anti-bonding σ* molecular orbital (MO) of the Cα-N bond in Asn263 (blue) interferes with the nascent bonding σ-MO of the C(=O) N-bond (red).

4.1.11 Cleavage does not affect the overall structure of wild-type YscU^C

Recombinantly produced wild-type YscU^C is readily cleaved into a ~ 6.3 kDa N-terminal (α-helices 0-1 and β-strand I) and a ~10.5 kDa C-terminal (α-helices 2-4 and β-strands II-V) fragment. The fragments, remain tightly linked, however, (Lavander *et al*., 2002) resulting in a single peak in gel-filtration chromatography (see Chapter 4.1.2). Cleavage of the wild-type protein thus does not cause significant overall structural rearrangements compared to the uncleaved YscU^C variants. This was confirmed by 2D-NMR ([^{15}N,^{1}H]-HMQC) spectroscopy by comparing ^{15}N-labelled wild-type and N263A YscU^C (Figure 29). Overall, wild-type YscU^C shows a good dispersion of the signals, characteristic for a folded globular conformation (Wüthrich, 1986). An equally good dispersion of the signals is observed for N263A YscU^C. The NMR-spectra of wild-type and N263A YscU^C superimpose very well, with most peak positions being virtually identical and line-widths apparently unaffected. Cleaved and uncleaved YscU thus share the same overall structure with minor, but physiologically highly relevant, changes in and close to the NPTH motif.

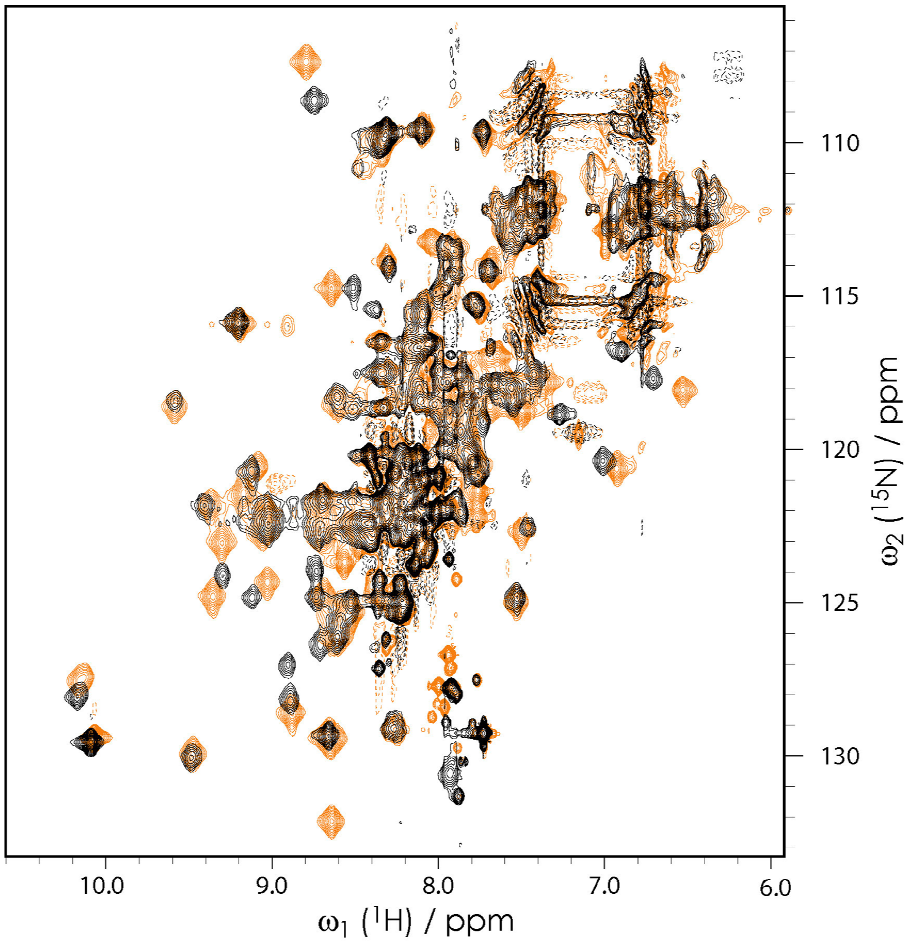


Figure 29: [15N,1H]-HMQC NMR spectra of YscU^{C}. The contour lines of wild-type YscU^{C} are shown in orange, those of YscU^{C} N263A in black.

4.1.12 The role of conserved amino acids in YscU^{C} cleavage *in vitro*

Substituting the nucleophile Asn263 by alanine, aspartate and glutamine abolished autocleavage of YscU. While Asn263 is thus essential for YscU-autocleavage, the role of surrounding residues is less clear. A systematic mutational analysis of the NPTH motif revealed an active participation of Pro264 and His266 in YscU autocleavage (Figure 30 A and B), whereas Thr265 is not involved (Sorg *et al*., 2007). Substitution of Pro264 by alanine reduced cleavage of YscU^{C} significantly, whereas substitution by glycine prevents cleavage altogether. Variant H266A showed a partial cleavage of YscU^{C}. The structure of YscU^{C} showed that two arginine residues (Arg296 and 314) flank the NPTH motif. Of these, Arg314 is highly conserved, while Arg296 is less conserved. Substituting Arg314 by alanine correspondingly reduced

cleavage partly whereas cleavage of variant R296A was unaffected. Interestingly, cleavage was entirely abrogated in the double variant H266A/R314A, suggesting a synergistic involvement of these residues in the mechanism of cleavage.

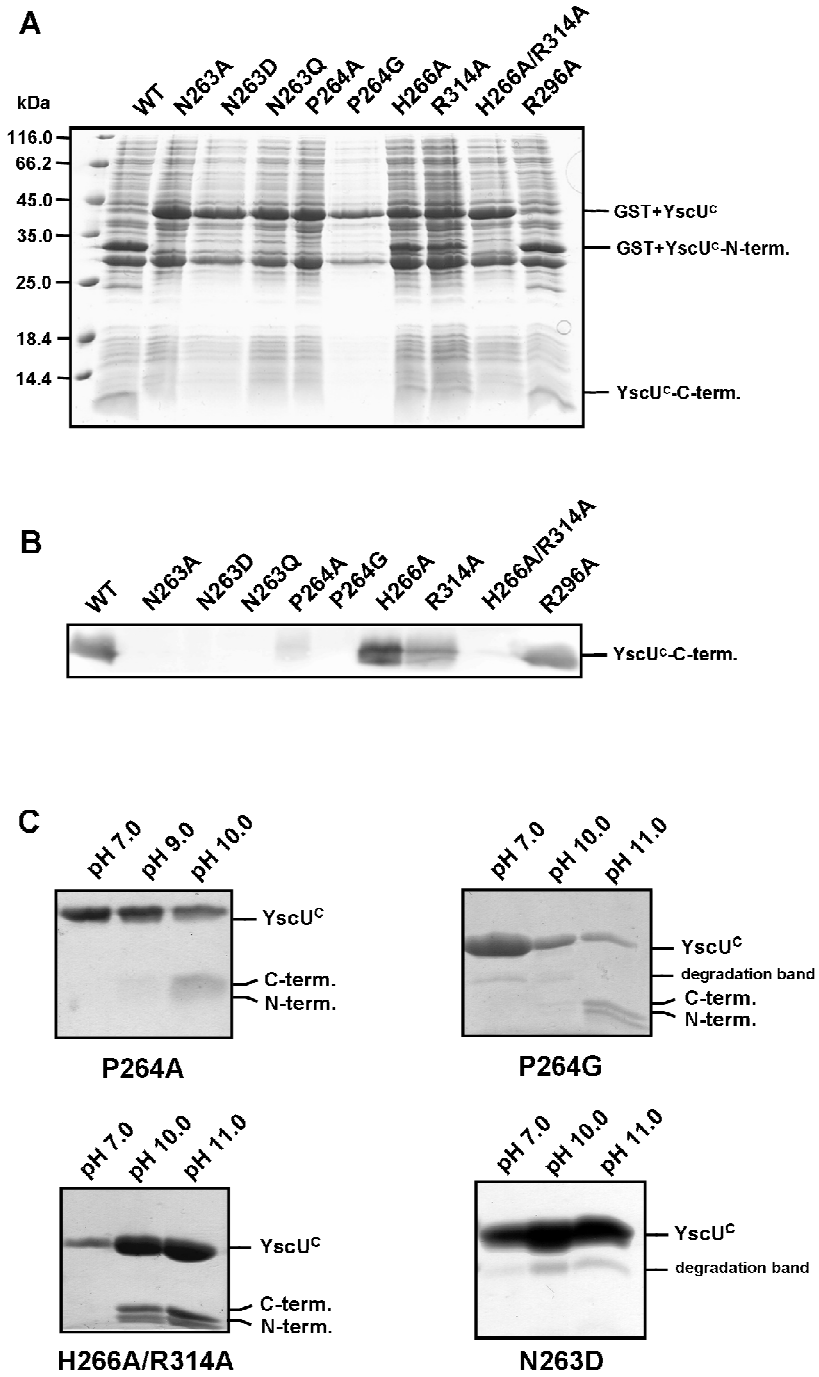


Figure 30: Autocleveage of different YscU^C variants expressed in E.coli. (A) Coomassie-stained 15% SDS-PAGE analysis of cells producing various YscU^C variants. (B) Overexpressed cells analyzed by immunoblotting using an anti-YscU^C antibody. (C) pH-dependent cleavage of YscU^C-variants. SDS-PAGE analysis of purified YscU^C-variants incubated at different pH-values.

4.1.13 Cleavage of some autocleavage-deficient YscU^C mutants is partially restored at high pH

Autocleavage of the slow-cleaving variant P270A of the flagellar protein FlhB^C (a homologue of YscU^C) has previously been reported to be pH-dependent (Ferris *et al*., 2005). YscU^C variants N263D, P264G, P264A and H266A/R314A were therefore analysed at pH 9, 10 and 11 (Figure 30 C). In variant N263D, no cleavage was observed at higher pH-values, whereas P264G, P264A and H266A/R314A were partially cleaved only at elevated pH. These observations match the published data on FlhB P270A (Ferris *et al*., 2005), suggesting that the nucleophile Asn263 needs to be activated by deprotonation of its side chain amide group.

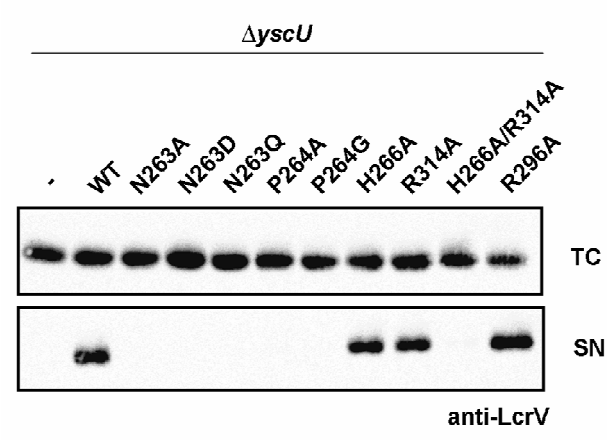


Figure 31 Analysis of LcrV secretion by different *yscU* mutants. Total cell (TC) and supernatant (SN) fractions of corresponding in trans complemented *Yersinia* Δ*yscU* mutant bacteria analyzed by immunoblot with anti-LcrV antibodies.

4.1.14 YscU mutants that are not cleaved *in vitro* do not export the LcrV translocator

Next, the export of the LcrV protein in the corresponding *Yersinia* mutant strains was investigated. Secretion was not affected for wild-type YscU and the variants H266A, R314A and R296A, but is lost in variants N263A, N263D, N263Q, P264A, P264G, and H266A/R314A, as confirmed by an immunoblot of the supernatant (Figure 31). The LcrV secretion pattern is in excellent agreement with the data described above for *in vitro* autocleavage.

4.1.15 The T3SS injectisome needle length is influenced strongly by YscU mutations

The influence of the YscU modifications on needle length was investigated by measuring needles from *Y. enterocolitica* Δ*yscU* mutant bacteria expressing the respective mutants *in trans*. Bacteria expressing wild-type *yscU* produced regulated needles of 69±10 nm. Similar needle lengths were measured for the YscU variants H266A (67±9nm), R314A (72±12nm), and R296A (76±11nm). In contrast, the length of needles in YscU variants N263A (125±40nm), N263D (143±51nm), P264A (112±42nm), and P264G (114±30nm) was increased and less regulated. Interestingly, YscU variants N263Q (80±16nm) and H266A/R314A (77±15nm) were regulated with a length distribution slightly larger than that of wild-type YscU (Figure 32).

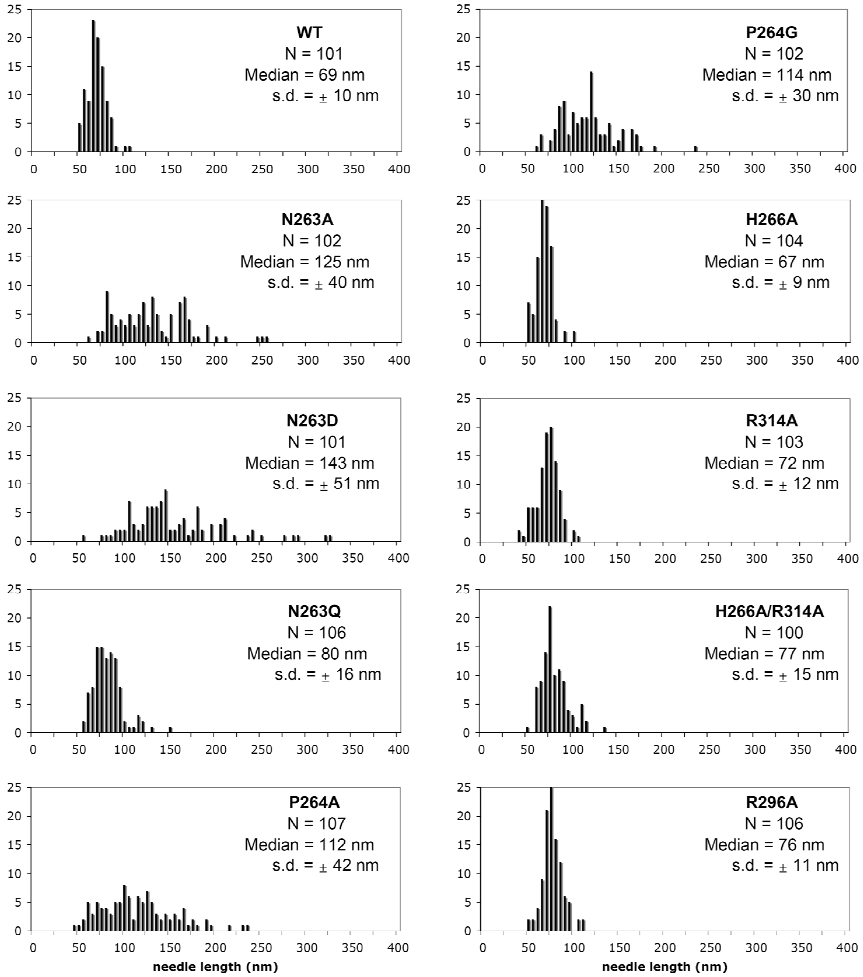


Figure 32: Needle length measurements of yscU mutant bacteria expressing the corresponding yscU alleles in trans from the pBAD promoter. Indicated are the median, the standard deviation (sd), and the number of needles measured (N).

4.2 The type III secretion apparatus protein YscV

All work in this thesis concerning the T3S apparatus protein YscV (formally called LcrD) was undertaken using the natural variant YscV R619G (wild type Swiss-Prot. Accession number P0C2V3). Hence in this thesis the natural variant YscV R619G is regarded as “wild type”.

4.2.1 Design of the YscV^{C} construct

Similar to YscU the T3S apparatus protein YscV can be divided into an N-terminal trans-membrane domain and a cytoplasmic C-terminal domain. An investigation of its primary sequence should reveal the potential domain borders suitable for creating an *E. coli* expression construct of the cytoplasmic domain. The analysis of the PHOBIUS plot (Figure 33) revealed six trans-membrane helices within the first 315 amino acid residues. A sharp increase in cytoplasmic probability at amino acid residue 322 indicates the start of the cytoplasmic domain. Therefore, the yscV^{C} gene (aa 322-704) was cloned into a pGEX-6P-1 vector using *Eco*RI and *Ava*I restriction sites. Gene expression results in a GST-YscV^{C} fusion protein, which is cleavable by the PreScission protease. For further experiments several variants of YscV^{C} were produced using site-directed mutagenesis.

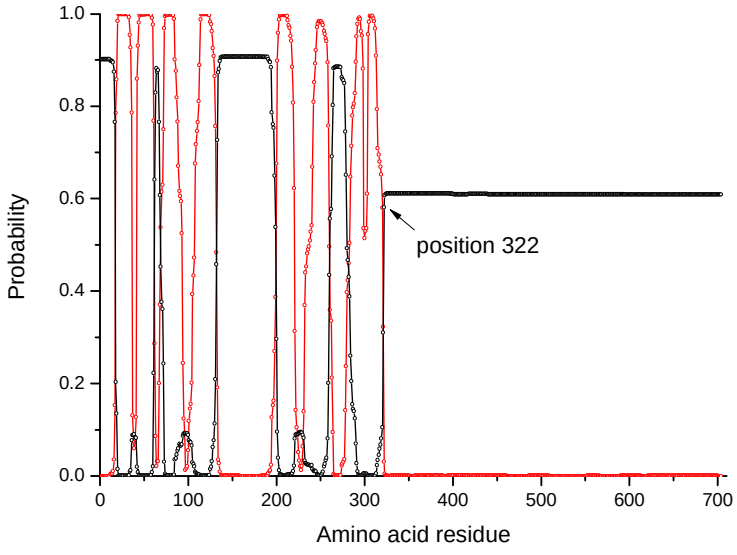


Figure 33: Topology and transmembrane plot of YscV. The plot shows the posterior probabilities of cytoplasmic (black) and TM helix (red) regions of YscV.

4.2.2 Preparation of YscV^C

A protein preparation scheme for YscV^C and its variants was established producing sufficient amounts of pure protein for further experiments. The following preparation scheme (Figure 34) was used for all YscV^C proteins.

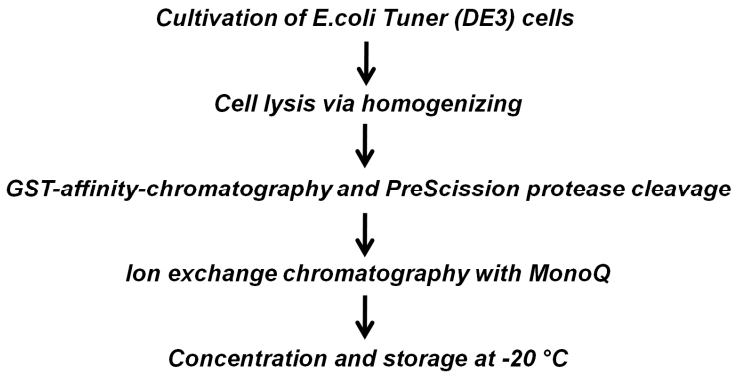


Figure 34: Protein preparation scheme for YscV^C and its variants.

Similarity to YscU^C GST-YscV^C was produced recombinantly in *E. coli* Tuner DE3 cells (see Chapter 3.3.1). After cell lysis supernatant was applied to glutathione sepharose for initial purification. Unbound protein was removed within three washing steps and GST-YscV^C was eluted with reduced glutathione in four steps (Figure 35). The GST fusion tag was cleaved off by adding PreScission protease. Cleavage was terminated by adding a protease inhibitor mix and subsequently applied to anion exchange chromatography. YscV^C was eluted in a single peak within a linear gradient (0.1 M -1 M NaCl) at 35 % ion exchange buffer B (Figure 36). The GST fusion protein was detected in the flow-through fractions and in a small peak prior to YscV^C main peak. Only weak upper bands were observed for the YscV^C containing fractions on the SDS-gel (Figure 36). Mass spectrometric analysis of the first upper band revealed that contamination is caused by the presence of the chaperone GroEL bound to YscV^C.

As the amount of contamination was minimal, the main peak fractions of YscV^C were pooled and concentrated to 10-15 mg/mL. Finally, about 5-10 mg protein was obtained from 1 L cell culture. A ThermoFluor assay (see Chapter 3.4.8) for YscV^C

revealed stabilizing conditions for YscV^{C} at 0.1 M Tris pH 8.1 and 100-200 mM NaCl (not shown). Therefore, protein solutions were dialysed against IEC buffer A and concentrated prior to crystallization.

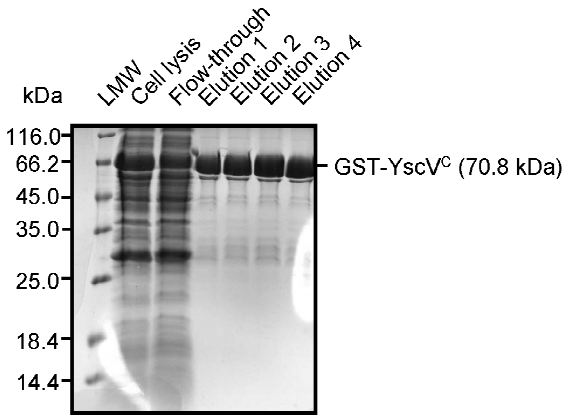


Figure 35: SDS-gel of the affinity chromatography of YscV^{C}.

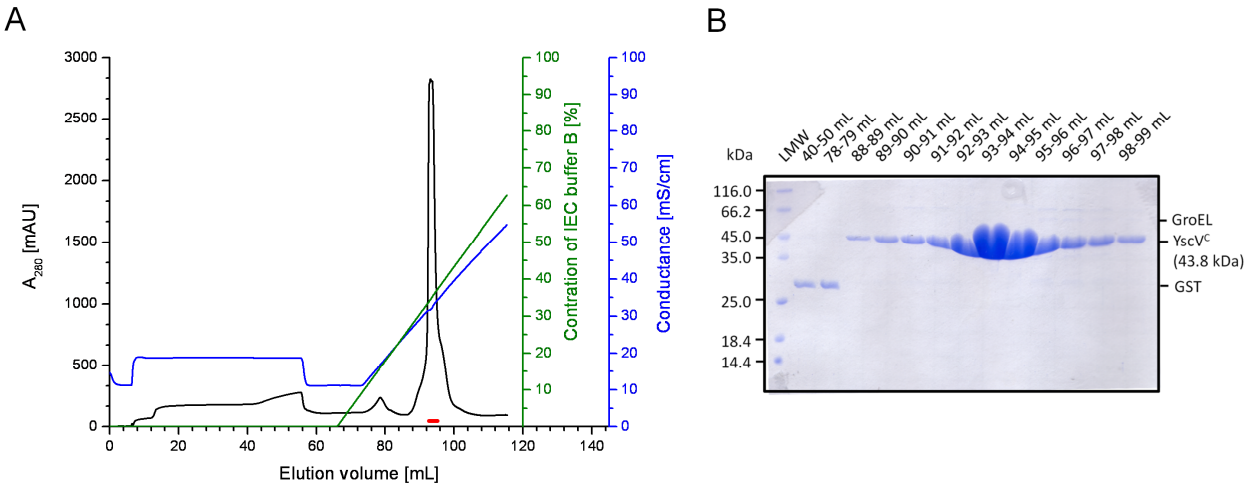


Figure 36: IEC elution profile (A) and corresponding SDS-gel (B) for YscV^{C}. Pooled fractions of YscV^{C} are depicted as a red bar in the elution profile.

4.2.3 Oligomerisation state of YscV^C is ambiguous

In order to get information on the oligomerisation state, YscV^C was analyzed by dynamic light scattering. To remove aggregated protein and dust particles samples were filtered and centrifuged prior to measurement. As YscV^C contains two cysteine residues, a second sample under reducing conditions (10 mM DTT) was also prepared. Table 9 shows a summary of the results obtained. Under both conditions the analysis revealed small amounts (up to 10 %) of aggregation, but the main peaks show an equal hydrodynamic radius indicating the absence of disulfide bonds. Assuming a globular shape of the protein, a potential molecular weight was calculated of about 116 kDa (YscV^C: ~ 44 kDa). Due to the low sensitivity of this method, YscV^C could therefore be a dimer or a trimer in solution.

To exclude any influence of the cysteine residues and to confirm the DLS results a test with Ellman´s reagent was performed (see Chapter 3.4.5). The negative test (A_{412}= 0.181) compared with YscV^C (A_{412}= 0.196) revealed only little increase of absorption, probably caused by the protein solution itself. This negative result confirms that both cysteine residues are not oxidized.

Table 9: Results of dynamic light scattering measurements.

Name	*R_{hyd} [nm]*	*Polydispersity [%]*	*MW [kDa]*	*Mass [%]*
YscV^C under oxidative conditions				
peak 1	4.6	3.1	119	94.0
peak 2	21.3	5.7	4339	5.3
YscV^C under reductive conditions				
peak 1	4.5	5.9	113	89.5
peak 2	17.2	9.4	2621	8.0

For further investigation of the oligomerisation state an analytical GPC run was performed (oxidizing conditions) using a Superdex S200 16/60 prep grade column. However, the elution profile shows multiple peaks containing YscV^C (Figure 37). The asymmetric peak (~ 48-60 mL) that is located close to the void volume indicates aggregated protein, which was also observed in the DLS experiment. The second

broad peak (~ 70 -96 mL) consists of at least two peaks with the same height, which could probably indicate dynamic changes in the oligomerisation state.

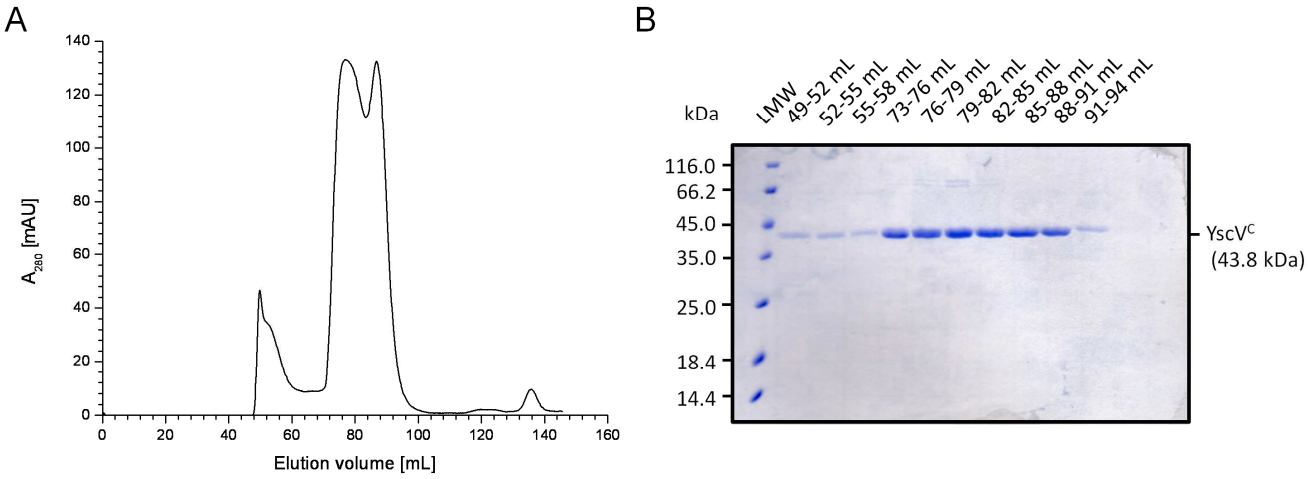


Figure 37: GPC elution profile (A) and corresponding SDS-gel (B) for YscV^C.

From the GPC run it was not evident whether YscV^C appeared as two distinct species or not. Therefore, an AF4 run with a coupled static light scattering experiment was performed (Figure 38). For this propose, volumes of 1.4 µL and 2.8 µL ($7x10^{-4}$ g/mol) were injected into the separation channel. In both runs YscV^C elutes in a broad peak similar to the GPC run. Analysis of the static light scattering experiment revealed increasing masses with progressing time. In the double concentrated run distinct plateaus are rising up at ~ 100, 126 and 210 kDa, which could be assigned to potential dimers, trimers and tetramers. However, the obtained results are only initial experiments indicating dynamic changes of several oligomerisation states. As YscV is supposed to be single species in the T3S apparatus, oligomerisation could be caused by the missing N-terminal domain. Therefore, further investigations of the oligomerisation state of YscV^C were not performed.

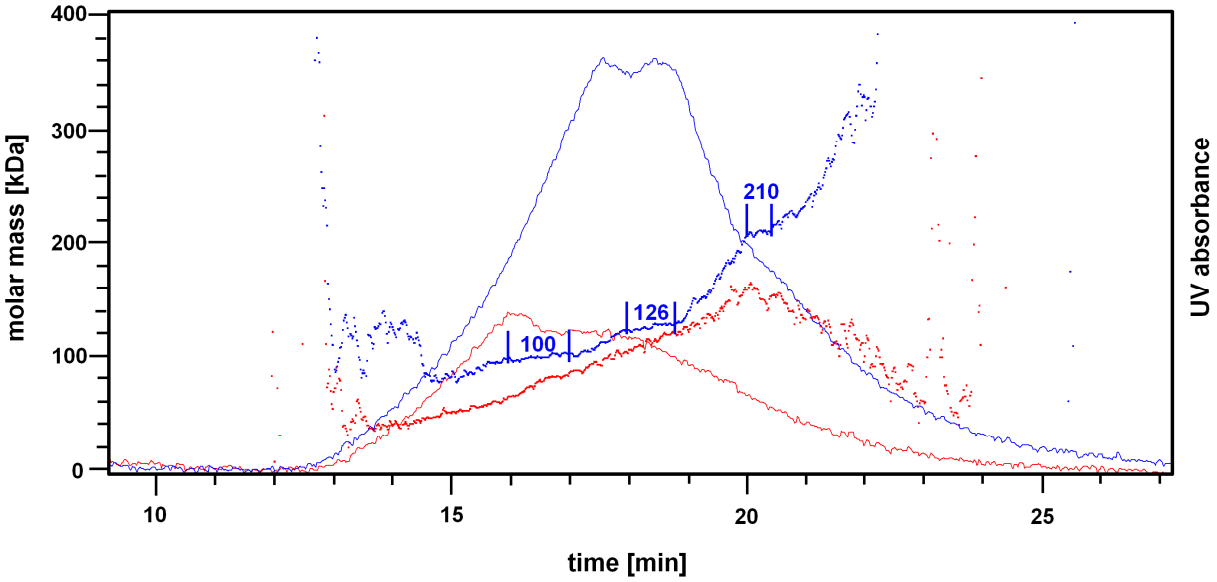


Figure 38: AF4 run of YscV^C with subsequent SLS analysis. Plot shows the UV absorbance (continuous line) and the molar mass (dashed line) as a function of time. Two runs were performed using 1.4 µL (red) and 2.8 µL (blue) YscV^C solution ($7x10^{-4}$ g/mol). The molar mass for distinct steps is measured as average.

4.2.4 Crystallization of YscV^C

Although the dynamic oligomerisation problem could not be resolved, initial screening experiments were performed using a crystallization robot. Drop sizes were set to 300 nL protein solution (10-15 mg/mL) and 300 nL reservoir solution. Screens were stored at 20 ℃ and controlled daily for protein cr ystals. After two days a shower of needle shaped crystals (Figure 39) was observed under several conditions (Table 10) containing either $(NH_4)_2SO_4$ or MPD (2-methyl-2, 4-pentandiol) as precipitation agent. Izit staining of the crystals confirmed the presence of protein in the crystals.

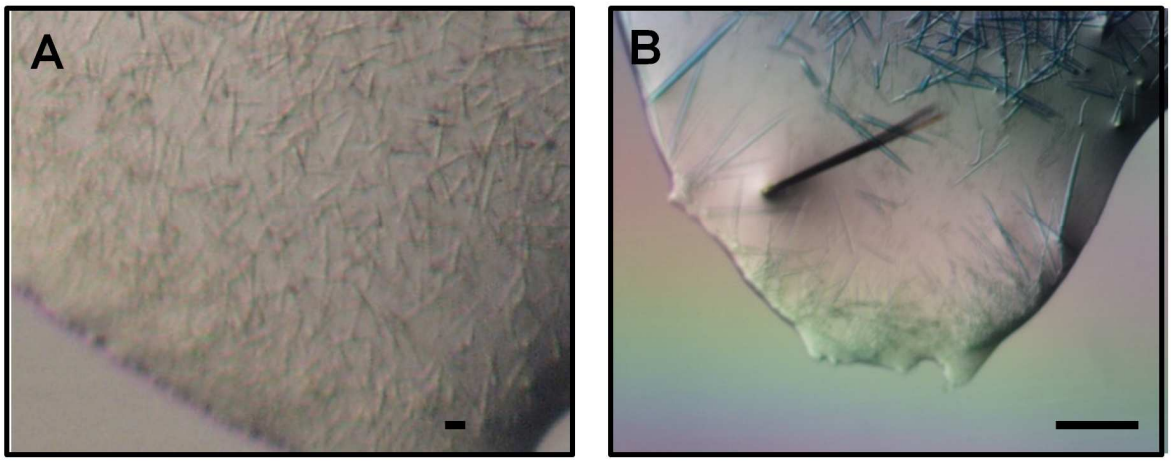


Figure 39: Initial crystals of YscV^C. Grown crystals under (A) NH_4SO_4 (JCSG Core suite III, E1) and (B) MPD (EasyExtal pHClear suite, G2) conditions as precipitation agent. In B crystals are stained with Izit. The scale bar is set to 300 µm.

Table 10: Initial crystallization conditions of YscV^C.

Screen	*No.*	*Condition*
EasyExtal AmSO4 suite	G5	0.1 M Tris pH 8.5, 1 M $(NH_4)_2SO_4$
EasyExtal AmSO4 suite	G6	0.05 M Sodium citrate, 1.2 M $(NH_4)_2SO_4$, 3 % (w/v) 2-propanol
EasyExtal pHClear suite	G2	0.1 M Sodium acetate pH 5.0, 10 % (v/v) MPD
EasyExtal pHClear suite	H2	0.1 M Sodium acetate pH 5.0, 40 % (v/v) MPD
EasyExtal pHClear suite	H8	0.1 M Sodium acetate pH 5.0, 65 % (v/v) MPD
JCSG Core suite II	C2	0.1 M Hepes pH 7.5, 2 % (v/v) PEG 400, 2 M $(NH_4)_2SO_4$
JCSG Core suite III	E1	0.1 M sodium cacodylate pH 6.0, 1.26 M $(NH_4)_2SO_4$
JCSG Core suite III	E2	0.01 M $CoCl_2$, 0.1 M Mes pH 6.5, 1.8 M $(NH_4)_2SO_4$
JCSG Core suite III	E3	0.1 M Mes pH6.5, 10 % (v/v) dioxane, 1.6 M $(NH_4)_2SO_4$

All crystallization conditions were tested for their reproducibility and crystal optimization was performed to yield crystals suitable for X-ray analysis. However, most of the crystals were difficult to reproduce and optimization for more suitable crystals failed. During the refinement process one condition (0.1 M Tris pH 8.5, 1 M $(NH_4)_2SO_4$ indicated the best for further refinement. Marcoseeding and additive screening yielded suitable crystals (Figure 40). However, the X-ray diffraction pattern

(BESSY BL14.1) range was limited to low resolution (R_{max}=~20 Å). In order to get better diffracting crystals co-crystallization of $YscV^C$ was tried with its potential binding partner $YscU^C$, but initial screening yielded only crystals of $YscV^C$.

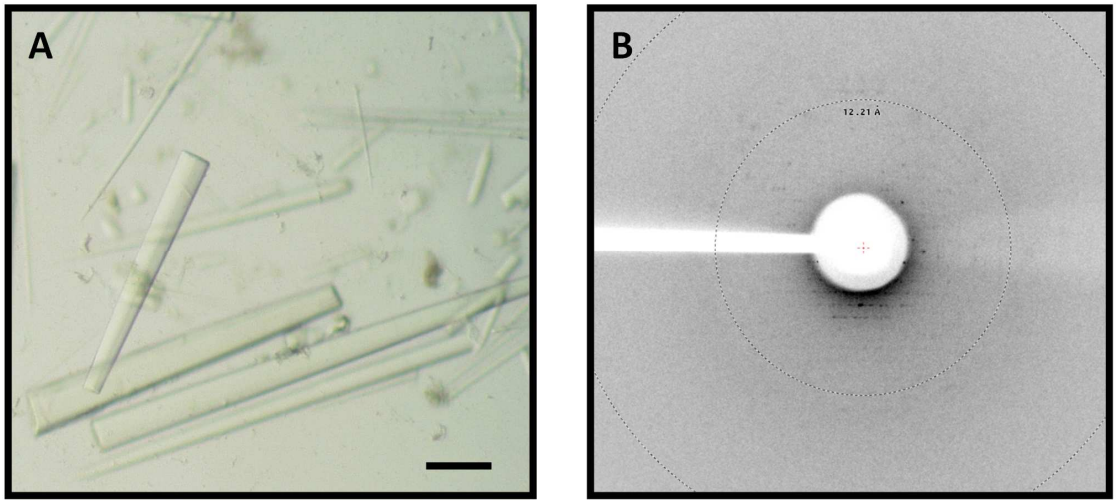


Figure 40: (A) $YscV^C$ crystals after macro seeding and (B) corresponding diffraction image. Scale bar is set to 100 Å. The lower resolution ring is set to 12.2 Å.

4.2.5 Limited proteolysis of $YscV^C$ created a new crystallization construct

As all crystallization attempts for suitable $YscV^C$ crystals failed, limited proteolysis was performed resulting a new crystallization construct. The flagellar protein homolog FlhA from *Salmonella thyphimurium* yielded well diffracting crystals (Saijo-Hamano *et al*., 2005) after limited proteolysis, as well. Therefore, a similar experiment was performed for $YscV^C$ using trypsin, α-chymotrypsin- and endoproteinase GluC. The protein solution (final concentration 0.3 mg/mL) was incubated with the different proteases (1:200, w/w) at room temperature. Samples were collected at defined times and the cleavage reaction was terminated by adding a mixture of SDS-sample buffer and AEBSF (0.1 M; 4-(2-Aminoethyl) benzenesulfonyl fluoride). For storage samples were frozen at -20℃.

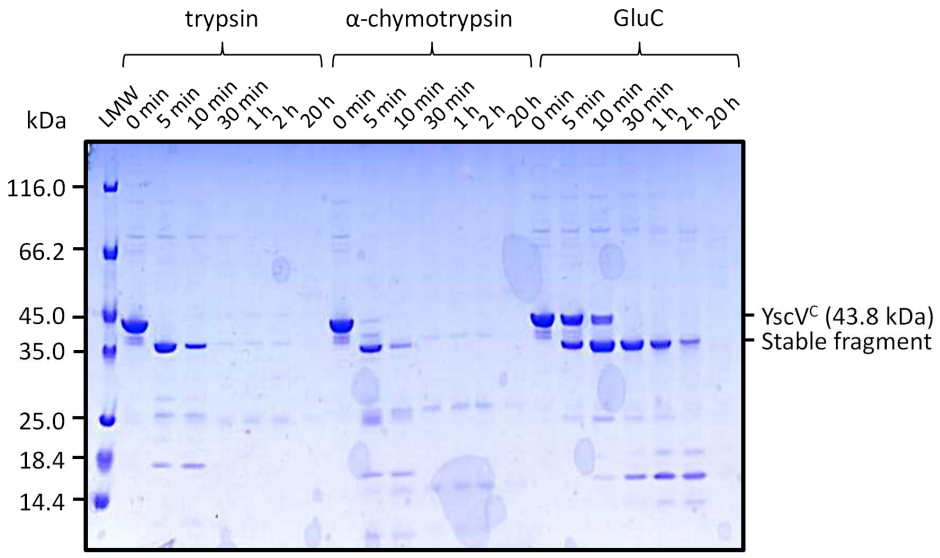


Figure 41: SDS-gel of a limited proteolysis of YscV^{C}.

Analysis of the limited proteolysis products (Figure 41) revealed an intensive band close to the 35 kDa marker for all three proteases. Using trypsin and α-chymotrypsin the band disappeared after 10 min, whereas with GluC the fragment was stable for over 2 h. Subsequent mass spectrometry analysis of a GluC gel band (10 min) revealed an intact C-terminus and an N-terminal fragment starting at amino acid residue 356. Further N-terminal sequencing analysis of the blotted trypsin digestion revealed the N-terminus starting at amino acid 360.

4.2.6 Protein preparation and crystallization of YscV$^{C}_{356}$

As a result of the limited proteolysis experiment, the new construct YscV$^{C}_{356}$ was cloned into a pGEX-6P-1 vector. Expression and purification were performed as described in Chapter 4.2.2. Analysis with DLS revealed a large R_{hyd} of 6.3 nm (polydispersity 15.3 %) and a test with Ellmann`s regent showed increasing absorption at 412 nm (negative test: A_{412}= 0.181, YscV$^{C}_{356}$: A_{412}= 0.364) indicating unspecific disulfide bonds. Therefore, a second dynamic light scattering experiment under reducing conditions (10 mM DTT) was performed and revealed a R_{hyd} (3.6 nm, polydispersity 17.8 %) which is the half R_{hyd} under oxidizing conditions. Further

preparations were then performed using 10 mM DTT as reducing agent. An analytical GPC run, which was performed under reducing conditions (10 mM DTT), showed a peak close to the void volume (44.4 mL) and a second one at 53.8 mL (Figure 42). The comparison of the elution volume to the molecular mass calibration curve (Figure 43) indicated a M_r of 480 kDa, which corresponds to a $YscV^C$ dodecamer assuming a globular shape.

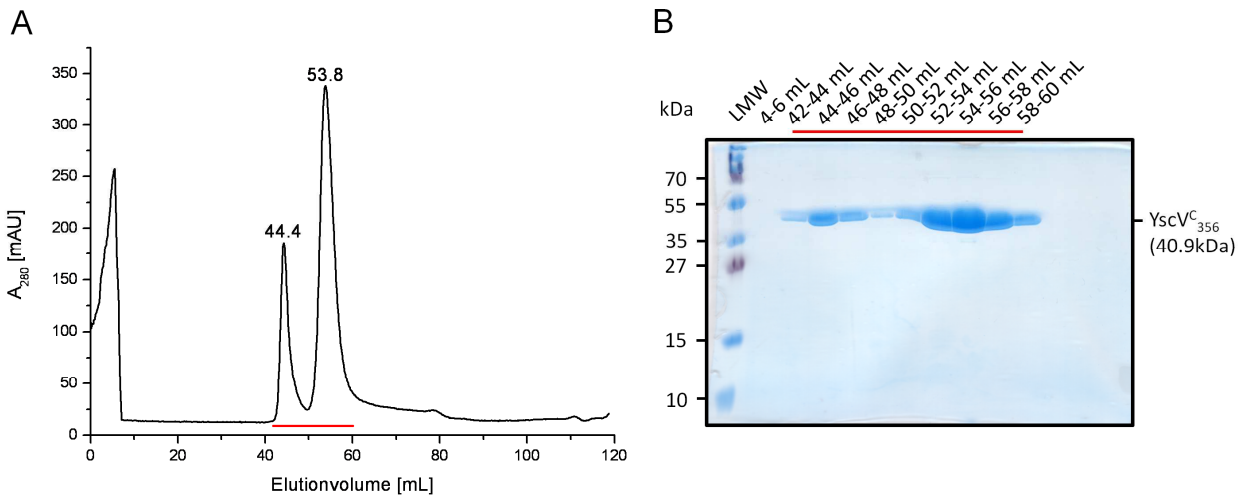


Figure 42: (A) GPC elution profil of $YscV^C_{356}$ on a Superdex 200 16/60 column and (B) corresponding SDS gel.

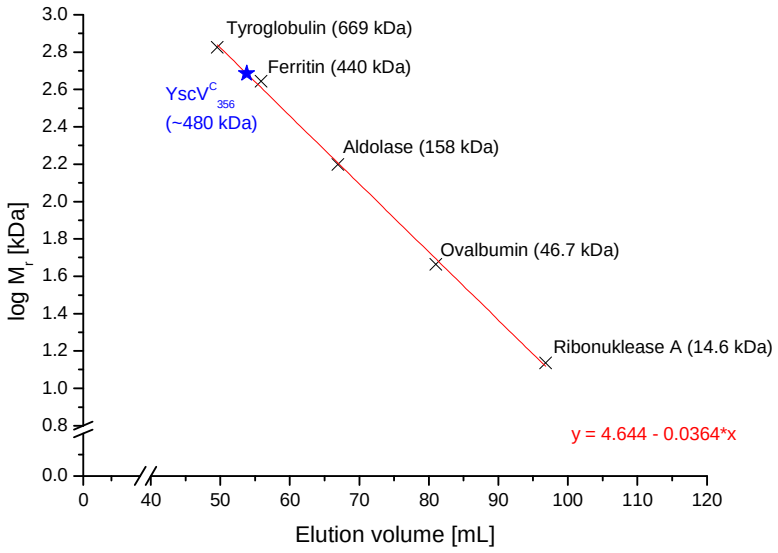


Figure 43: Calibration curve of Superdex S200 16/60 and linear fit. The calibration run was performed using four different calibration proteins and a flow rate of 0.6 mL/min.

For initial crystallization screening the YscV$^{C}_{356}$ protein solution was reduced with buffered TCEP (*tris*(2-carboxyethyl)phosphine; final concentration 5 mM), as TECP is more resistant to oxidation in air as other reducing agents. Droplets were used with 200 nL reservoir solution and 200 nL protein solution (24 and 12 mg/mL). After 2 d long needle shaped crystals (Figure 44) were obtained under several conditions (Table 11). X-ray suitable crystals were harvested from their mother liquid (SM I suite B12) and flash frozen in liquid nitrogen, as cryo protection was not necessary. However, X-ray diffraction patterns were not observed at our home-source. A mass spectrometric analysis of dissolved crystals was performed confirming YscV$^{C}_{356}$ to be present in the crystals (Figure 45). This was achvied by dissolved of crystals (washed before) in a 50 % (v/v) glycerol solution. Due to time constraints further crystal improvement and refinement has not yet been undertaken.

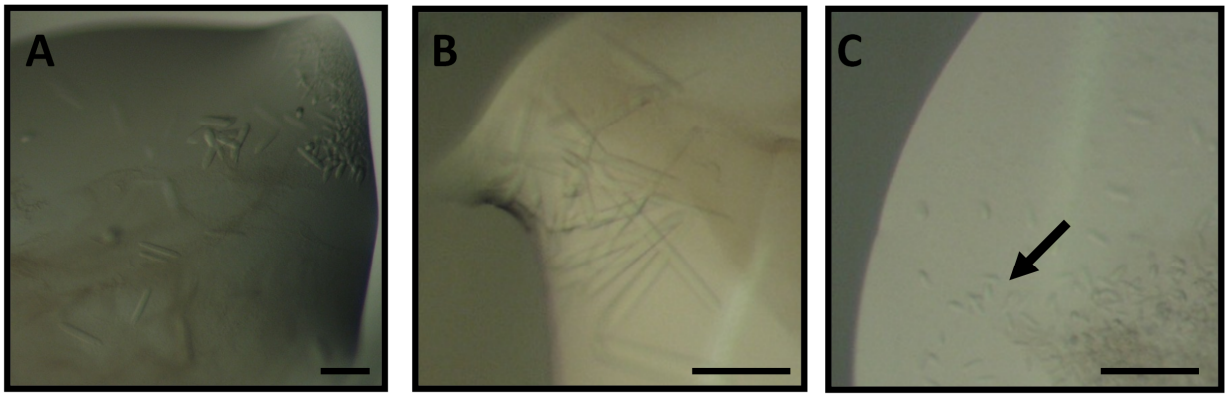


Figure 44: Initial crystals of YscV$^{C}_{356}$ under reducing conditions. (A) SM I suite, B12; **(B)** SM I suite C1; **(C)** JCSG Core suite III, D1. The scal bar is set to 100 µm.

Table 11: Initial crystallization conditions for $YscV^{C}_{356}$.

Screen	*No.*	*Condition*
SM I suite	B12	0.2 M NaCl; 0.1 M imidazole pH 8.0; 1 M $(NH_4)_2HPO_4$
SM I suite	C1	0.1 M sodium actate pH 4.5; 1 M $(NH_4)_2HPO_4$
SM I suite	G8	0.2 M Li_2SO_4*H_2O; 0.1 M Ches pH 9.5; 1 M K/Na-tartrate*4H_2O
SM I suite	G11	0.2 M NaCl; 0.1 M imidazole pH 8.0; 1 M K/Na-tartrate*4H_2O
JCSG Core suite III	D1	1 M K/Na-tartrate; 0.2 M NaCl; 0.1 M imidazole pH 8.0
JCSG Core suite III	E4	1.6 M $MgSO_4$; 0.1 M Mes pH 6.5
JCSG Core suite IV	G2	1 M K/Na-Tartrate; 0.1 M Mes pH 6.5
JCSG Core suite III	E2	0.01 M $CoCl_2$, 0.1 M Mes pH 6.5, 1.8 M $(NH_4)_2SO_4$
JCSG Core suite III	E3	0.1 M Mes pH6.5, 10 % (v/v) dioxane, 1.6 M $(NH_4)_2SO_4$

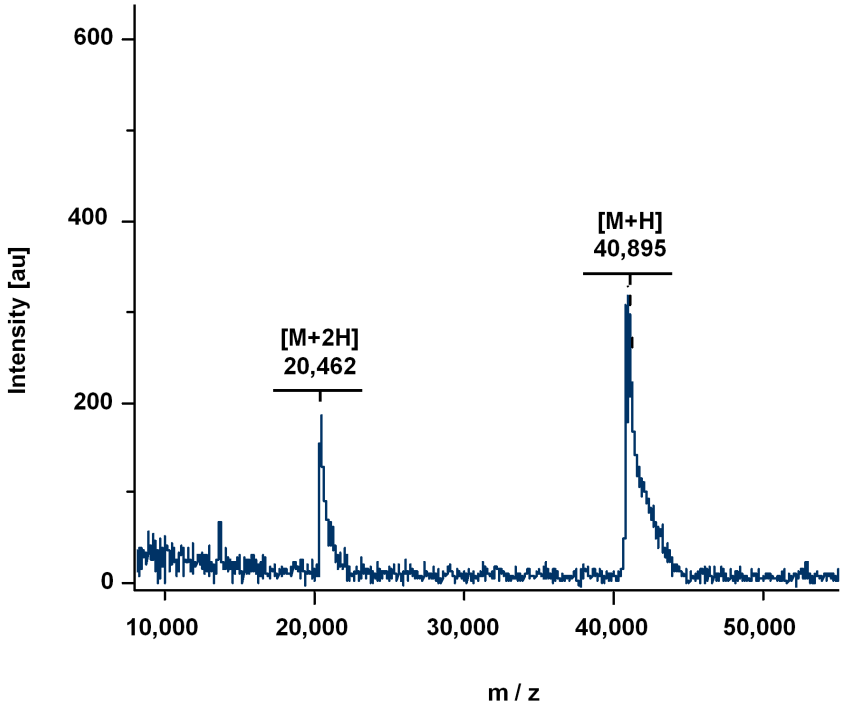


Figure 45: MALDI-TOF-MS spectrum of a dissolved crystal of $YscV^{C}_{356}$. [M+H] indicates the singly-charged species, [M+2H] the doubly-charged ion. The calculated molecular mass of $YscV^{C}_{356}$ (including additional amino acid residues that remain on the protein after GST cleavage: GPLGSPEF) is 40,912 kDa.

4.3 Interaction analysis between YscU^C and YscV^C

Analysis of the flagellar export apparatus proteins FlhA (corresponding to YscV) and FlhB (YscU) revealed interactions for full-length proteins and their cytoplasmic regions (Minamino & MacNab, 2000b; Zhu *et al*., 2002). Further interaction analysis by FRET measurement suggested an interaction for the latter, which is dominated by an electrostatic and/or hydrogen-bonding network (Zhu *et al*., 2002). To analyze the nature of this interaction in the *Yersinia* system, epitope scans were performed between YscU^C and YscV^C. For this purpose, corresponding 15mer peptides were produced and spotted on a cellulose membrane. The membranes were prepared and incubated with the corresponding proteins and antibodies. Used membranes were recycled for further experiments or evacuated in a shrink wrap film and stored at -20 ℃.

In the first analysis round interactions were identified and interaction sites were roughly located. Therefore two membranes were produced with peptide spots containing short sequences of YscU^C and YscV^C. The peptide spots were arranged as 15mers, which were shifted by three amino acid residues along the amino acid sequence (15/3 scan). To exclude unspecific interaction between peptide spots and GST, a background scan with GST was performed (Figure 46). The YscV^C membrane revealed unspecific interactions for five peptide spots, whereas no unspecific interaction was detected for YscU^C. All other membranes also showed no unspecific interactions.

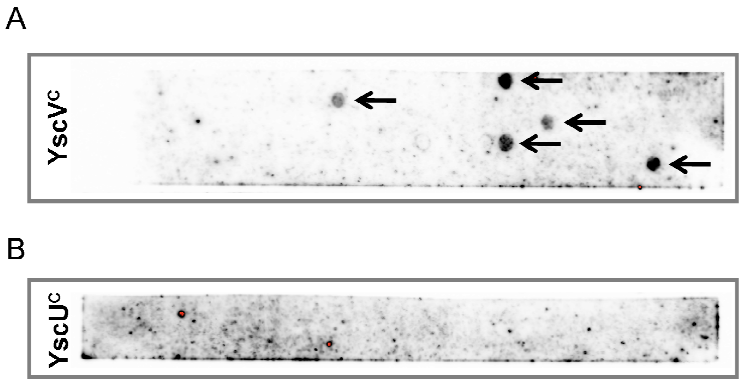


Figure 46: Background scans for YscV^C (A) and YscU^C (B). Membranes were incubated with GST to detected unspecific interactions. Peptide spots with unspecific interactions are marked by black arrows.

Only unspecific background spots and no specific interactions were observed between GST-YscU^{C} and the YscV^{C} peptides (Figure 47 A). However, the YscU^{C} peptides (Figure 47 B), which were incubated with GST-YscV^{C}, revealed interaction spots for peptides 18-21, 24-25, 32-34 and 43. The scan showed interaction spots with weak and stronger intensity providing information about binding affinity between peptide and protein. For precise analysis of the binding sites, areas around intensive spots (18-25 and 31-35) were scanned again using a 15/1 epitope scan. A finer 15/1 scan for peptide 43 was not necessary, as the peptide is located at the C-terminal end of YscU^{C} and the spot before showed no interaction at all. Furthermore, the intensive peptide spots 21, 24, 33 and 43 were analyzed in an alanine epitope scan revealing the binding potential of each amino acid residue between peptide and protein.

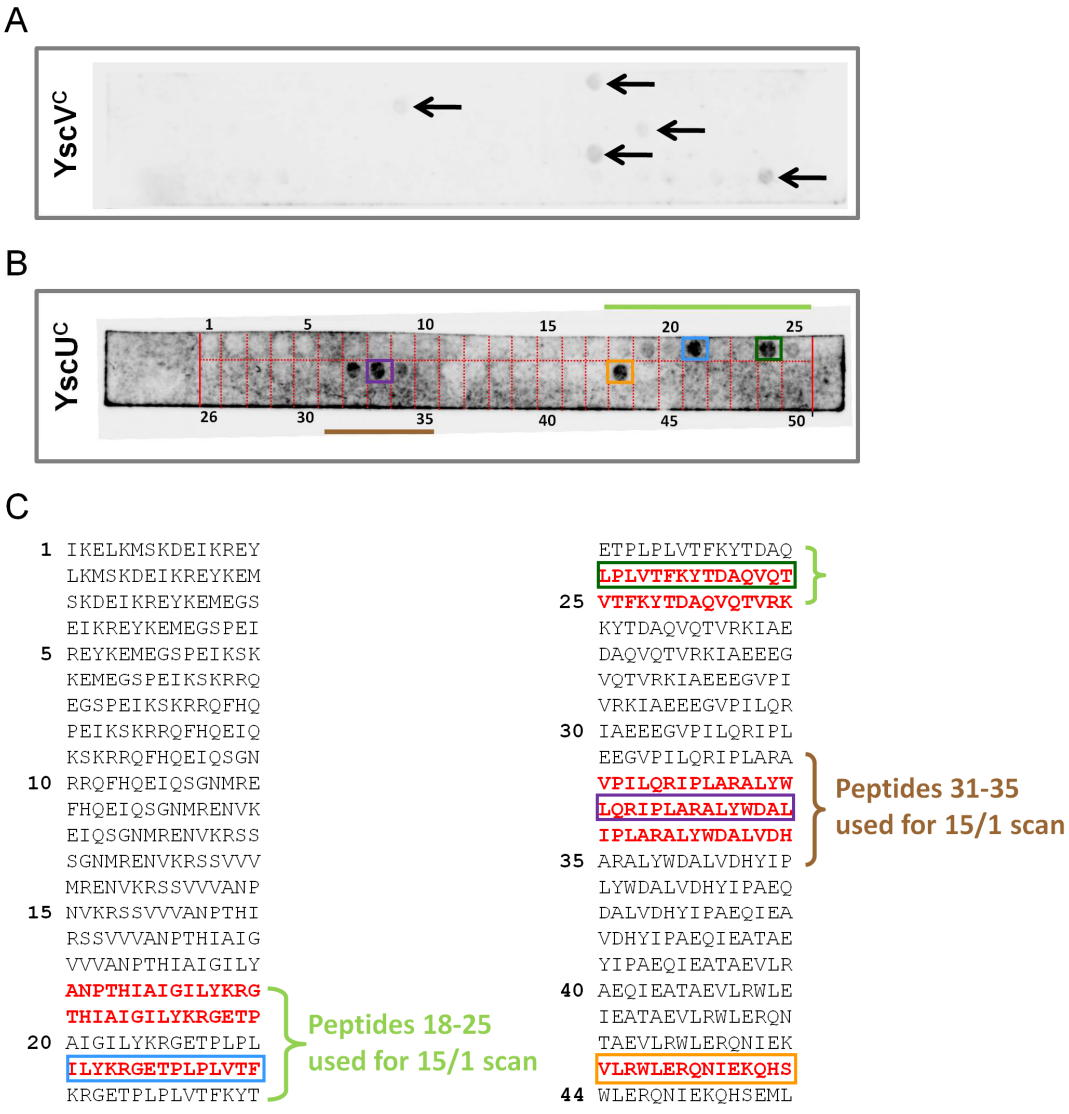


Figure 47: 15/3 epitope scans of YscV^C (A) Unspecific interactions are marked with black arrows. **15/3 epitope scans of YscU^C (B) and corresponding sequence analysis (C).** Interacting peptide sequences are highlighted in red. Peptides used for a 15/1 scan and an alanine scan are indicated with a coloured bar and square, respectively.

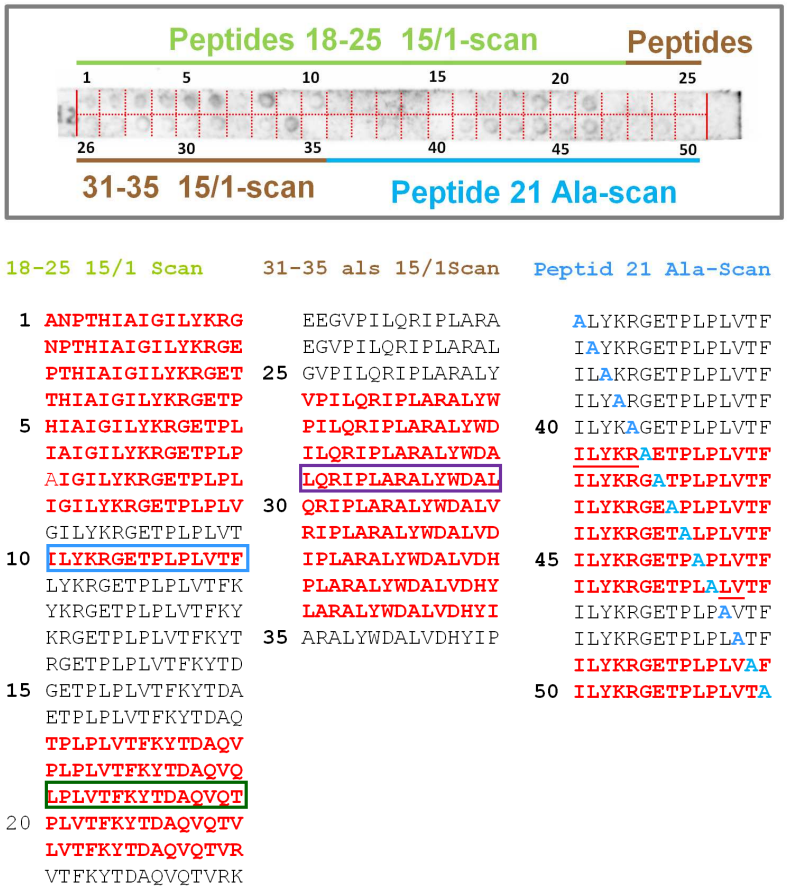
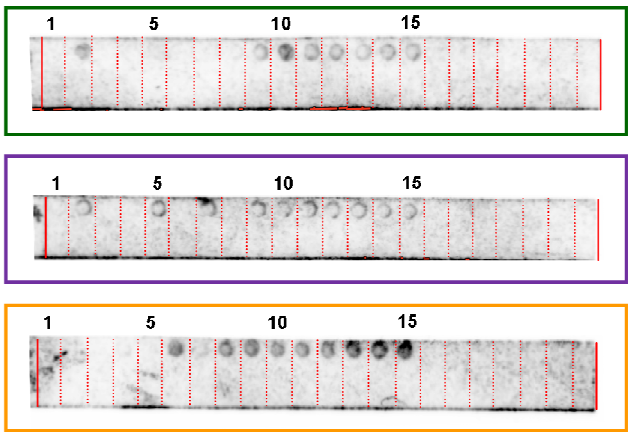


Figure 48: 15/1 and Ala-scans of YscU^{C} peptides. Interacting peptide sequences are highlighted in red. Peptides used for a 15/1 scan and an alanine scan are indicated with a coloured bar and square, respectively.

The analysis of the 15/1 epitope scan for peptides 18-25 (first round, light green bar) showed two main interaction areas including peptides 1-10 and 17-21 (Figure 48). Within these areas peptides share two common sequence sites: $_{271}$ILYKRG$_{276}$ (1-10) and $_{281}$LVTFKYTDAQV$_{292}$ (17-21). A further alanine scan analysis (Figure 48 and Figure 49) revealed important binding sequences as $_{271}$ILYKRx$_{276}$ (1-10) and $_{281}$LVTFKYxxxx$_{292}$ (17-21) within the motifs. The second 15/1 scan for peptides 31-35 (first round, brown bar) showed interactions with peptides 26-34 (Figure 48) with the common sequence site $_{312}$LARALYW$_{318}$ and the alanine scan provided $_{312}$LxRxxx$_{318}$ (Figure 49) as important amino acid residues. For the last peptide 43 (orange) only an alanine scan was performed revealing the binding motif $_{337}$VLRWLxR$_{343}$ (Figure 49).

Finally, the epitope scan interaction analysis yielded four different peptide interaction sites for GST-YscU^{C}. Within these sites an alanine scan revealed essential amino acids for binding of peptides. Transfer of the epitope scan results on the structure of YscU^{C} are discussed in the next chapter.



Peptid 24 Ala-Scan	Peptid 33 Ala-Scan	Peptid 43 Ala-Scan
1 APLVTFKYTDAQVQT	1 AQRIPLARALYWDAL	1 ALRWLERQNIEKQHS
LALVTFKYTDAQVQT	LARIPLARALYWDAL	VARWLERQNIEKQHS
LPAVTFKYTDAQVQT	LQAIPLARALYWDAL	VLAWLERQNIEKQHS
LPLATFKYTDAQVQT	LQRAPLARALYWDAL	VLRALERQNIEKQHS
5 LPLVAFKYTDAQVQT	5 LQRIALARALYWDAL	5 VLRWAERQNIEKQHS
LPLVTAKYTDAQVQT	LQRIPAARALYWDAL	VLRWLARQNIEKQHS
LPLVTFAYTDAQVQT	LQRIPLARALYWDAL	VLRWLEAQNIEKQHS
LPLVTFKATDAQVQT	LQRIPLAAALYWDAL	VLRWLERANIEKQHS
LPLVTFKYADAQVQT	LQRIPLARALYWDAL	VLRWLERQAIEKQHS
10 LPLVTFKYTAAQVQT	10 LQRIPLARAAYWDAL	10 VLRWLERQNAEKQHS
LPLVTFKYTDAQVQT	LQRIPLARALAWDAL	VLRWLERQNIAKQHS
LPLVTFKYTDAAVQT	LQRIPLARALYADAL	VLRWLERQNIEAQHS
LPLVTFKYTDAQAQT	LQRIPLARALYWAAL	VLRWLERQNIEKAHS
LPLVTFKYTDAQVAT	LQRIPLARALYWDAL	VLRWLERQNIEKQAS
15 LPLVTFKYTDAQVQA	15 LQRIPLARALYWDAA	15 VLRWLERQNIEKQHA

Figure 49: Ala-scans of YscU^{C} peptides. Interacting peptide sequences are highlighted in red. Alanine exchanged amino acid residues, which showed no interaction, are underlined.

5 Discussion

5.1 YscU - a type III secretion recognition protein

YscU is one of five integral membrane proteins (YscR, -S, -T, -U, -V) building the T3SS export apparatus, which is likely to be located in the inner membrane ring of the basal body (Cornelis, 2006). Its autocleavable cytoplasmic domain plays a critical role in recognition of substrates during the export process (Sorg *et al*., 2007). To gain structural information about the T3S recognition protein YscU, the autocleaving cytoplasmic domain was intensely analyzed. For this purpose, a preparation scheme for $YscU^C$ was established yielding sufficient amounts of protein. Crystal structures of $YscU^C$ variants N263A and N263D were determined using multiple anomalous dispersion. The high resolution structure revealed the structural basis of $YscU^C$ autocleavage. Additionally, single-residue $YscU^C$ variants within the NPTH motif or its immediate vicinity were investigated with respect to their role in cleavage and needle-length control.

5.1.1 YscU is poised for autocatalytic cleavage

The crystal structure of non-cleaving $YscU^C$ variant N263D (the asparagine nucleophile is replaced by the isosteric homologue aspartate) suggests that the side-chain of Asn263 is positioned ideally for a nucleophilic attack and subsequent cleavage of wild-type $YscU^C$. In N263D, the attack on its own carbonyl is prevented by the insufficient nucleophilicity of the Asp 263 carboxylate, despite its equivalent position. In N263Q, the nucleophilicity should be similar to that of asparagine. However, the longer and hence more flexible side chain disfavours the formation of a lactam ring in this variant, preventing cleavage of $YscU^C$.

$YscU^C$ autocleavage is crucially dependent on the positioning of the attacked carbonyl group of Asn263, which is part of the peptide bond to Pro264. Increasing the flexibility of the main chain correspondingly prevents cleavage in the variants P264A and P264G. As indicated above, cleavage is compromized in variants H266A and R314A, but entirely abrogated in H266A/R314A. Increasing the pH increases

cleavage, suggesting a role for both His266 and Arg314 in efficient autocleavage. By providing of a hydrogen bond, Arg314 most likely serves to optimally position the side chain of Asn263 allowing nucleophilic attack to occur at an angle of 109°. The role of His266 is, however, less clear. It may potentially mediate the abstraction of a proton from Asn263 either directly or by deprotonating bridging water molecule. A similar role has been proposed for a histidine residue during autosplicing of the DnaB mini-intein (Ding *et al*., 2003). Thr265, as the only not fully conserved amino acid of the NPTH motif, does not appear to be involved directly in the cleavage reaction (Sorg *et al*., 2007; Zarivach *et al*., 2008). A preference for β-branched amino acids indicates a role in structurally stabilizing the type II turn (Gunasekaran *et al*., 1997)

The results are supported by similar structural and functional data reported for the YscU homologues EscU from enteropatogenic *E. coli*, SpaS from *Salmonella typhimurium* (Zarivach *et al*., 2008), Spa40 from *Shigella Flexneri* (Deane *et al.*, 2008) and YscU from *Yersinia pestis* (Lountos *et al*., 2009). The crystal structures of the cleaved wild type EscU, as well as a number of EscU variants, were determined and correlated with *in vivo* cleavage experiments of full-length EscU (variants). Wild type $YscU_{pestis}$, EscU, SpaS and Spa40 reveal unchanged tertiary structures of nicked proteins, concurring with our conclusion based on NMR-analysis of $YscU^{C}$ wild type and mutant (Figure 50).

The only major differences concern $YscU^{C}$ variants H266A and R314A and their homologous variants *in vivo* in full-length EscU. Both $YscU^{C}$ variants show partial autocleavage *in vitro* and a wild type-like secretion of LcrV *in vivo*. EscU variant H265A is found to be fully autocleaved but secretion of T3S substrates is lost (Zarivach *et al*., 2008). EscU variant R313A, by contrast, remains uncleaved but again T3S substrates secretion is lost (Zarivach *et al*., 2008). A third EscU variant, R133T, is fully cleaved but retains normal effector secretion (Zarivach *et al*., 2008). Thus, despite a high level of structural conservation between YscU and EscU, some functional differences between the *Yersina* and EPEC T3SS may exist. The *in vitro* cleavage data on $YscU^{C}$ His266 variants would presumably implicate His265 in EscU in the asparagine-driven autocleavage reaction.

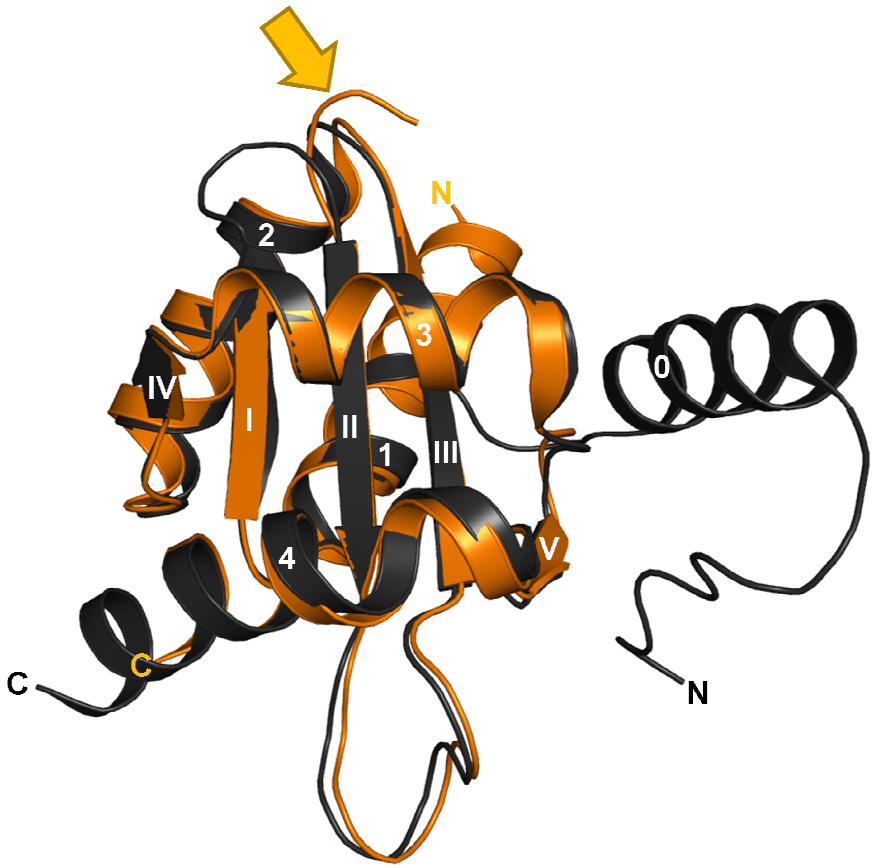


Figure 50: Structural overlay of YscU^{C}N263D (2W0R) and YscU$^{C}_{pestis}$ (2JLI). The cleaved N'PTH-motif is depicted by an arrow.

Based on structural and functional data the following autocleavage reaction mechanism is proposed (Figure 51). (i) The unique reaction geometry of YscU and its homologues is a prerequisite for the nucleophilic attack. The β-amide nitrogen, activated through proton abstraction by His266 or an intermediate water molecule, is optimally positioned to attack the electrophilic carbonyl carbon at an angle of 109° (108° in EscU). (ii) The tetrahedral reaction inter mediate is stabilized by hydrogen bonds to the backbone amides of His266 and Ile267. (iii) Protein cleavage gives rise to a labile succinimide, which is rapidly hydrolysed to yield a new C-terminal asparagine or iso-asparagine.

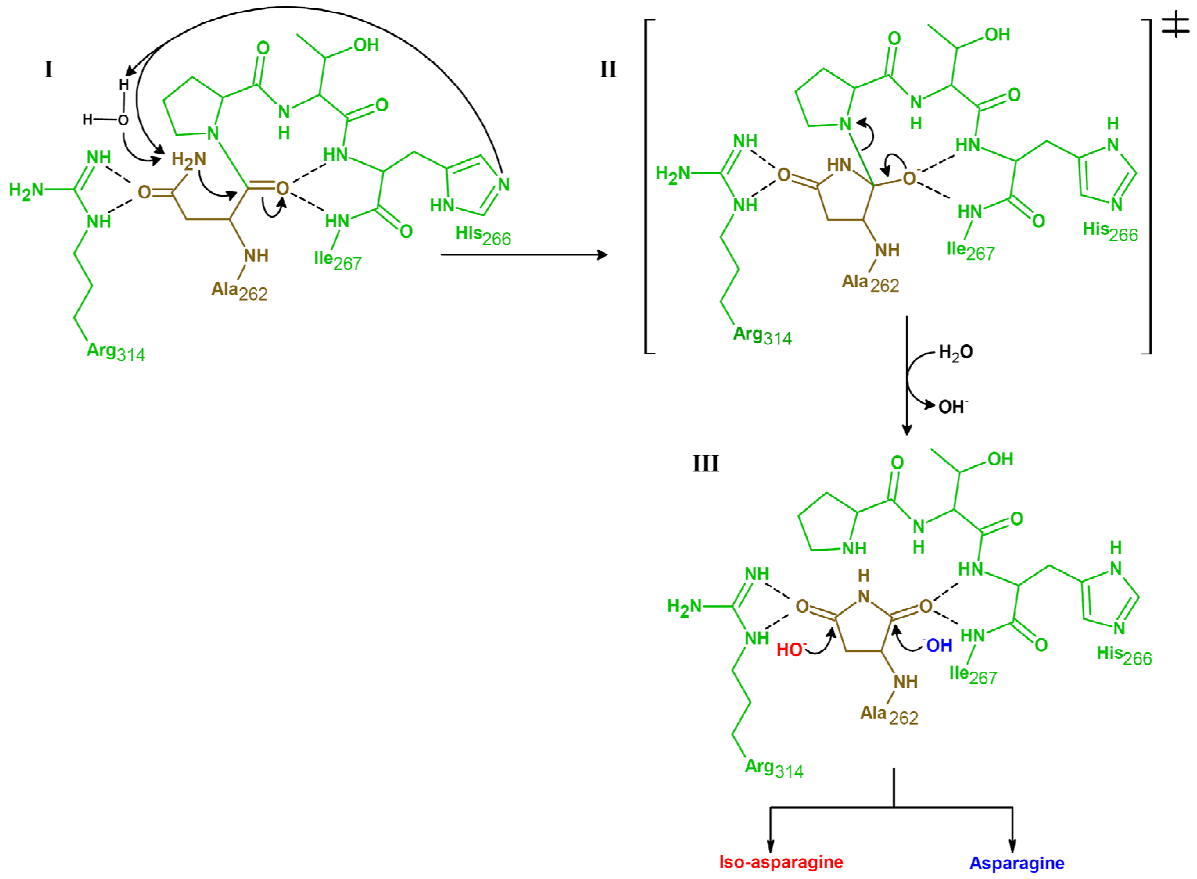


Figure 51: Structure-based model of succinimide-mediated cleavage of YscU. (I) His266 (or a His266-activated water molecule) abstracts a proton from the β-amide nitrogen of Asn262. The latter nucleophilically attacks the Asn262 carbonyl resulting in the rate-limiting tetrahedral intermediate (II). Cleavage of the peptide bond yields a labile C-terminal succinimide (III) which is hydrolysed to asparagine or iso-asparagine.

The structure of YscU^{C} also explains phenotypes of previously reported YscU variants. The strongly reduced export of the presumed inner rod protein YscI (in an *yscPyscU* mutant background) in the YscU variant Y317D implicates Tyr317 in interacting directly with YscI (Wood *et al*., 2008). Structurally, this strictly conserved residue is fully exposed at the surface of α-helix 3. Presumably, the introduction of a negative charge prevents interaction with YscI explaining the observed phenotype (Zarivach *et al*., 2008). The YscU variant G270N has been reported to cause the complete loss of T3S and YscU autocleavage (Riordan & Schneewind, 2008). Gly270, in the middle in β-strand II, is fully buried in the hydrophobic core of the domain. Replacing this residue by the significantly bulkier asparagine presumably

compromises folding or structure of the protein, resulting in degradation of this protein and consequently, a non-functional T3SS.

5.1.2 Needle length in a non-cleavable YscU mutant is controlled even without YscP over-expression

In vivo analyses reveal a new phenotype for the YscU variants N263Q and H266A/R314A. These variants do not undergo autocleavage *in vitro*. Corresponding bacterial mutants do not secrete the translocator LcrV that constitute the tip complex thus prevent pore formation. Their needles are, however, of normal size, although significantly less molecular ruler protein YscP is secreted compared to wild type bacteria (data not shown). In the YscU N263A mutant, wild type needle length could be restored only by YscP over-expression implying that YscU autocleavage inhibition diminishes YscP secretion losing control over needle length (Sorg *et al*., 2007). The phenotype observed here for the variants N263Q and H266A/R314A demonstrates that the loss of autocleavage may indeed affect the secretion of YscP and LcrV without relinquishing control over needle length. The fact that different phenotypes are observed for N263A and N263Q involving the same residue, may indicate that small amounts of YscP are sufficient to determine needle length but that a defined conformation of YscU also contributes.

5.1.3 The flexible linker region is composed of two hinged α-helices

A structural superposition of uncleaved $YscU^C$ variant N263A with $EscU^C$ (rmsd 1.3 Å for 94 common Cα atoms), $SpaS^C$ form *Salmonella typhimurium* (rmsd 1.0 Å for 80 Cα atoms), and $Spa40^C$ from *Shigella flexneri* (rmsd 1.0 Å for 82 Cα atoms) reveals virtually identical folds for the cytoplasmic core region and the NPTH motifs, despite a low level of overall sequence identity (Figure 52).

An obvious difference between $YscU^C$ and its homologues $EscU^C$, $SpaS^C$ and $Spa40^C$ involves the proposed flexible linker (211-248) connecting the globular autocleavage domain to the N-terminal membrane-spanning four-helix bundle (Figure

52). In YscU^{C}, amino acids 232-244 form a well-defined α-helix, which is not observed in the homologues structures except for uncleaved YscU$^{C}_{pestis}$ (Lountos *et al*., 2009). In the crystal structure of YscU^{C}, α-helix α0 is involved in a crystal contact to α-helix α4 of the next monomer. One of the crystal forms of EscU^{C} additionally reveals an additionally α-helix (α-1) N-terminal of α0 (Zarivach *et al*., 2008). The presence of α-helix α0 confirms earlier predictions of the secondary structure for the flagellar protein FlhB (Fraser *et al*., 2003).

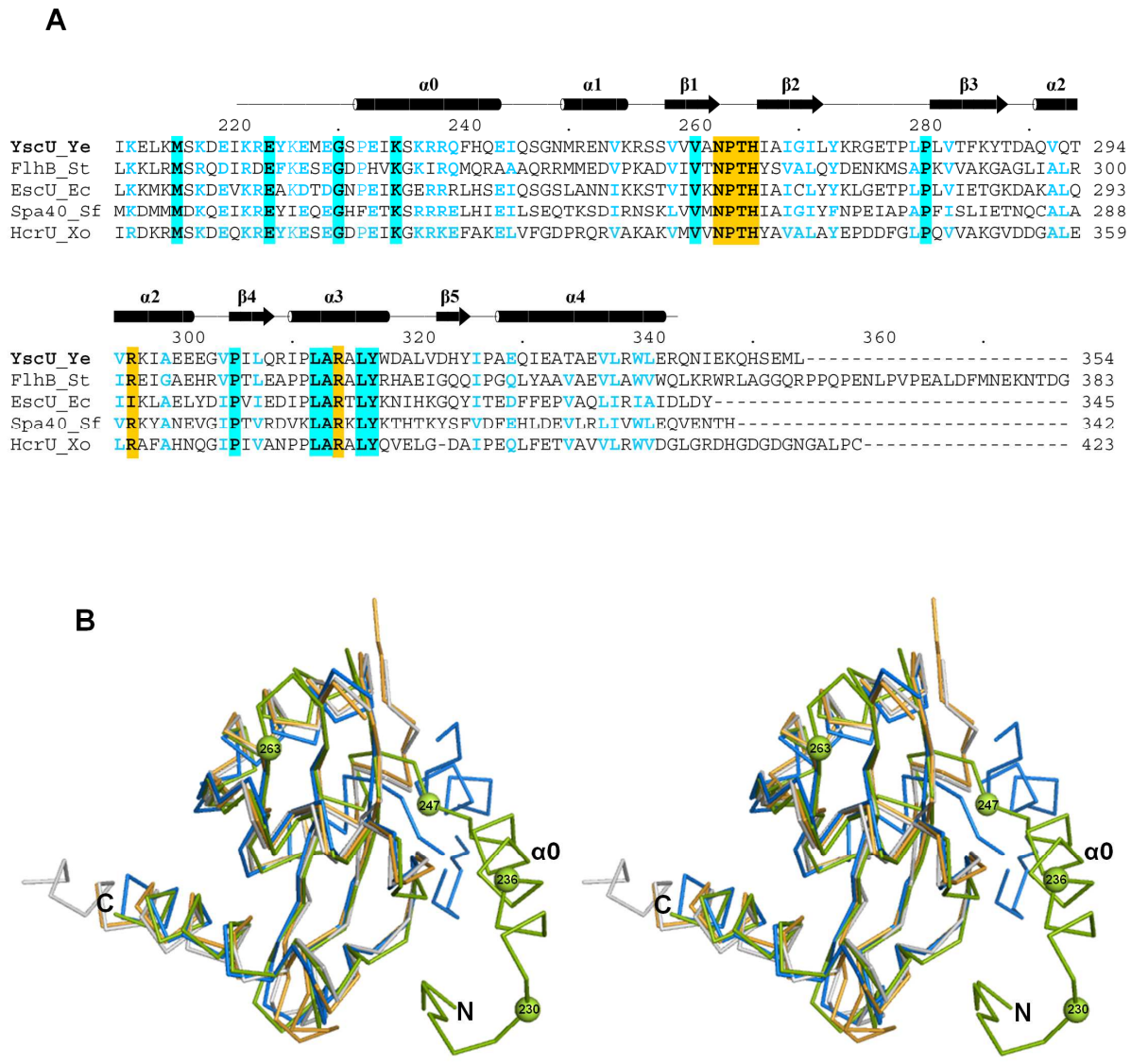


Figure 52: (a) Structure-based sequence alignment of YscU^{C} from Y. enterocolitica (Swiss-Prot. accession number Q56844) with FlhB^{C} from S. thyphimurium (P40727), EscU^{C} from E. coli (Q7DB59), SpaS^{C} from S. thyphimurium (P40702), Spa40^{C} from S. flexneri (Q6XVW1) and HcrU^{C} from X. oryzae (Q5H6T1). The secondary structures are from YscU^{C} N263D. Conserved amino acids are marked in bold letters on coloured backgrounds. Substituted residues in this study are indicated by a yellow background. Less conserved amino acids are also bold and coloured in blue.

(b) Stereoview $YscU^C$ variant N263D (green) superimposed on its wild-type homologues $EscU^C$ (blue), $SpaS^C$ (grey) and $Spa40^C$ (orange). Important positions are marked by a green sphere.

The linker region is essential for T3S (Zarivach *et al*., 2008). Corresponding deletions 232-236, 234-245 and 230-245 in EscU and the point mutation G235P (S236P in YscU) abolish the export of T3S substrates, despite intact autocleavage. Similar results have been obtained for the FlhB protein (Fraser *et al*., 2003). Mapping these mutations onto the structure of $YscU^C$ indicates that each would delete or kink α-helix α0, which presumably functions as a spacer element, separating the globular autocleavage domain from the inner membrane. Additionally, the helix could be involved in direct or indirect binding of other components.

Small amino acids such as Gly229 and Gly248 in EscU are assumed to provide flexibility to the linker region (Zarivach *et al*., 2008). Substituting conserved Gly229 by proline in EscU (G230 in YscU and located N-terminal of α-helix α0) decreases the export of T3S substrates significantly. An equivalent substitution in Gly247 (G248P in YscU and located C-terminal of α0) also reduces the export of T3S substrates. Combining the structural information on $YscU^C$ and $EscU^C$ with the *in vivo* data suggests that YscU-like proteins are composed of three domains (Figure 55): A membrane-spanning N-terminal domain and a C-terminal, cytoplasmic, globular, autocleavage domain bridged by a flexible linker region that can switch from an extended to a partially α-helical conformation. This adaptable architecture of the linker allows for sufficient conformational freedom to enable interactions of the autocleavage domain with other components of the T3SS during its assembly.

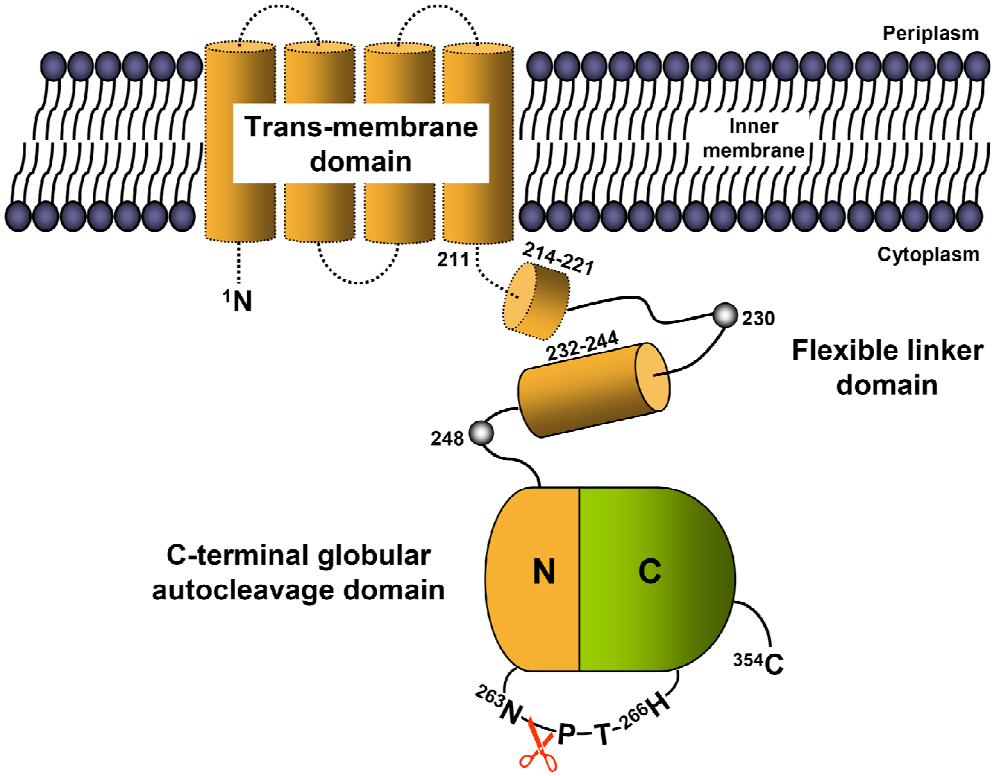


Figure 53: Model of tripartite domain arrangement of YscU: The N-terminal, membrane-spanning domain is followed by a flexible linker region and a globular autocleavage domain.

5.2 The inner membrane protein YscV

The integral membrane protein YscV (~ 77.8 kDa) is the largest component within the T3S apparatus. It is composed of an N-terminal trans-membrane domain, which is predicted to contain at least six trans-membrane helices, and a ~ 43 kDa C-terminal cytoplasmic domain. For the flagellar homolog protein FlhA interaction has been shown between its cytoplasmic domain and the ATPase complex as well as the cytoplasmic domain of YscU (McMurry *et al*., 2004; Zhu *et al*., 2002). Thus, it has been proposed that YscV and YscU together build a platform for docking the ATPase-substrate complex (Minamino *et al.*, 2010).

For structural characterization $YscV^{C}$ was produced as a GST fusion protein and a purification scheme was established yielding sufficient amounts of the cytoplasmic domain. However, the obtained protein was slightly contaminated with GroEL, a

chaperone that is required for proper protein folding. Consequently, it can be assumed that small amounts of $YscV^{C}$ are not well folded. Investigation of the oligomerization state by DLS and analytical GPC revealed a multimeric oligomerisation state that is not induced by disulfide bridges. Further analyses with AF4 and SLS indicate a dynamic oligomerisation (up to tetramers). In contrast to these results the flagellar homolog $FlhA^{C}$ was clearly observed as a monomer (Saijo-Hamano *et al.*, 2004). Besides, it has been proposed that export apparatus proteins have only a low copy number (one or maximal two) (Minamino *et al.*, 2008; Van Arnam *et al.*, 2004). Therefore the high oligomerisation state is probably the consequence of not properly folded protein due to the missing N-terminal part.

Despite the tendency for oligomerisation, screening for $YscV^{C}$ protein crystals revealed several crystallization conditions that contain either MPD or $(NH_4)_2SO_4$ as precipitant. However, X-ray suitable crystals showed only diffraction patterns within low resolution range. The obtained spots are narrowly spaced indicating huge cell dimensions and therefore probably less ordered crystals.

Analog to FlhA, a limited proteolysis experiment revealed a shorter N-terminus confirming a linker region between the trans-membrane domain and a compact cytosolic domain (Minamino *et al.*, 2010; Zhu *et al.*, 2002). The resulting truncated variant $YscV^{C}_{356}$ has to be treated with reducing agents, as aggregation was observed under oxidizing conditions. $YscV^{C}_{356}$ seems to be well folded, as no chaperone contamination was observed. However, not a monomer but a single, highly oligomeric species (~480 kDa) was obtained as analyzed by GPC. Crystallization trials revealed several initial hits containing needle shaped crystals. However, X-ray investigations revealed no X-ray diffraction. Due to time constraints, a detailed crystal refinement could not be performed. But these results provide a good basis for further crystal improvement.

5.3 Interaction analysis between YscU^C and YscV^C

An interaction between the flagellar cytoplasmic domains of FlhA and FlhB has been shown by affinity blotting and FRET analysis to be dominated by electrostatic and/or hydrogen-bonding within the potential interaction site (Zhu *et al*., 2002). In order to characterize the interaction between YscV^C and YscU^C, a peptide-based epitope scan was performed. Thereby, the cellulose anchored peptides mimic surface areas of the corresponding interaction partner.

Initial experiments revealed interactions of YscU^C peptides with YscV^C protein, but not the other way around. This is probably due to the fact that amino acid residues, which are involved in the interaction, are often close to each other in the folded protein, which is not true for elongated sequences. Therefore, the epitope scan was not suitable to observe complex three dimensional interactions. In a second analysis round the located interaction sites were analyzed in detail by a 15/1 scan resulting in four possible binding peptides: $_{271}$ILYKRG$_{276}$, $_{281}$LVTFKYTDAQV$_{292}$, $_{312}$LARALYW$_{318}$ and $_{337}$VLRWLER$_{343}$.

By mapping the results on the structure of YscU^C N263A, four distinct interacting regions (Figure 54) can be assigned: tip of β-strand 2 and the following loop, β-strand 3 and the following loop and α-helices 3 and 4. All regions are located within the C-terminal half of YscU^C. The amino acid residues are accessible for solvents except β-strand 3, which is mainly buried by α-helices 0. It can be assumed that the main part of β-strand 3 is solvent accessible, although α-helix 0 covers this site in the crystal structure. As it was shown (see Chapter 4.1.9), the linker is pinched in this position due to crystallization contacts. The alanine scan provided information of side chain contributions to the interaction between peptide and protein. Analysis of the side chain accessibility revealed several amino acid side chains, which are important for interaction in the epitope scan, but are buried in the crystal structure. This mismatching concerns mainly amino acids of β-strand 3 and the amino acids Ile271 and Leu312. The differences indicate the limitation of this method, mimicking the surface of a β-strand. However, several complex structures show that peptides are able to form helical interaction surfaces (Okon *et al*., 2008). Therefore, one can conclude that the highly conserved α-helix 3 and the end of α-helix 4 are important

for binding YscV^{C} to YscU^{C}. However, this presumption has to be proved by further interaction experiments as well as docking experiments based on structural information of both potential binding partners.

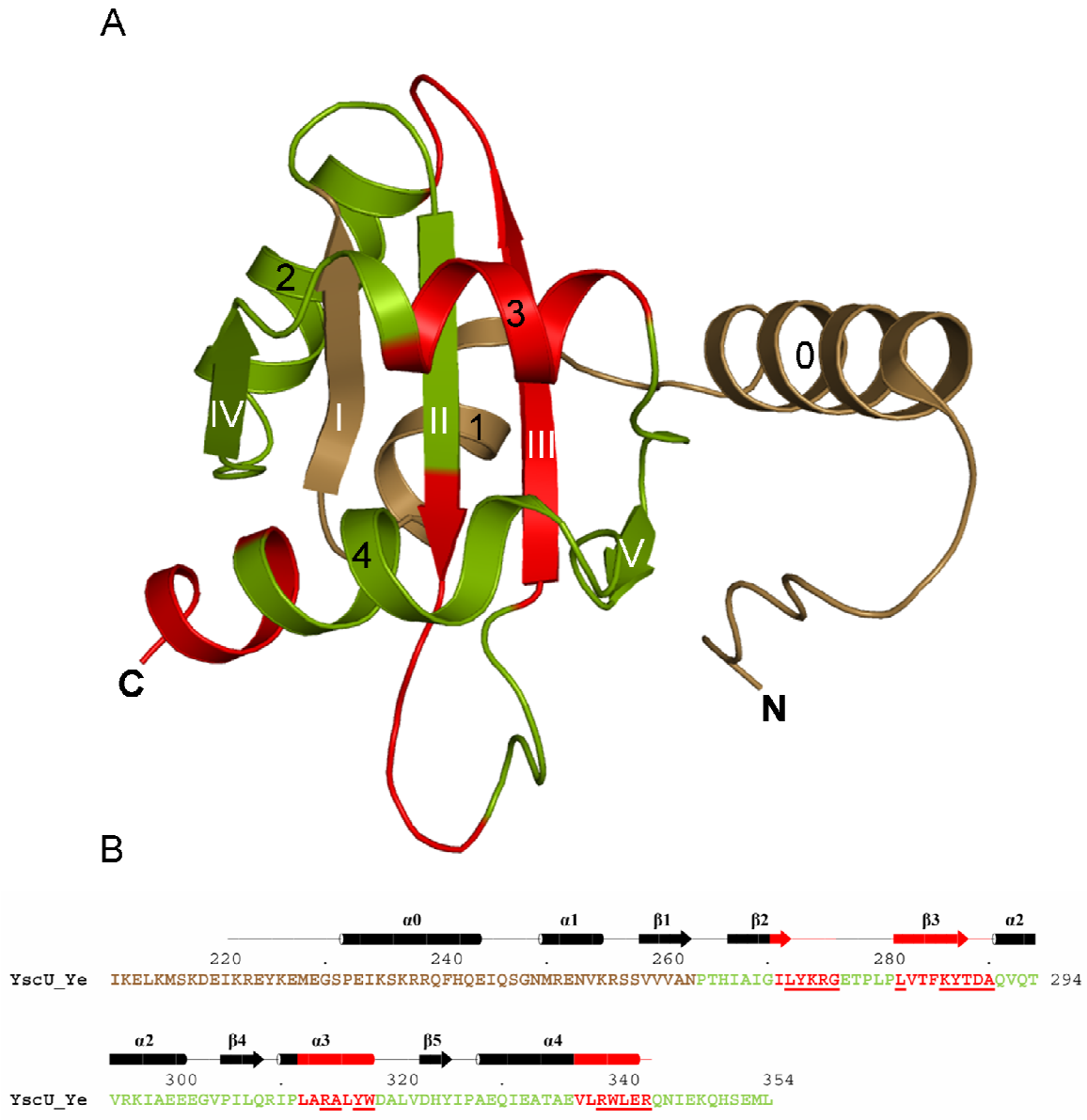


Figure 54: Proposed interaction regions between YscU^{C} and YscV^{C}. (A) Tertiary and **(B)** secondary structure of YscU^{C} N263A. The proposed binding regions are colored in red. Side chains of underlined amino acid residues are accessible for solvents. However, analysis of the alanine epitope scan revealed several side chains as important for interaction, which are buried in the crystal structure ($_{271}$ILYKRx$_{276}$, $_{281}$LVTFKYxxxx$_{292}$, $_{312}$LxRxxx$_{318}$ and $_{337}$VLRWLxR$_{343}$, underlined aa side chains are not solvent accessible).

6 Outlook

Recently, several crystal structures of substrate specificity proteins belonging to the T3SS YscU protein family have been determined in its cleaved and uncleaved form (Deane *et al.*, 2008; Lountos *et al.*, 2009; Wiesand *et al.*, 2009; Zarivach *et al.*, 2008). The structural basis for autocleavage has been elucidated by analyzing several non-cleavable mutants. All data confirm that autocleavage occurs immediately after recombinant expression in *E. coli* cells. However, natively purified full-length YscU showed also uncleaved YscU suggesting that the cleavage process is regulated (Sorg *et al*., 2007). For the flagellar system it has been proposed that autocleavage is regulated in a time-dependent manner by the protein itself, further proteins or external factors such as pH or inhibitory proteins. It still remains elusive how the cleavage event is regulated (Deane *et al.*, 2010). Further, the structures reveal a linker region between trans-membrane and autocleavage domains. The proper function of the linker region is essential for secretion of both effectors and translocators. Nevertheless, it is unclear how the linker is involved in the secretion process.

Besides functional questions concerning the structural parts, YscU has also been shown to be involved in several interactions including the ATPase regulator complex (YscK/L) and the proposed C-ring (YscQ) (Riordan & Schneewind, 2008). Further experiments have to be conducted to analyze the role of these interactions during secretion. Furthermore, it has been proposed that YscU and YscP together regulate the needle length by triggering the so-called “substrate specificity switch” (Edqvist *et al*., 2003). Investigations of the *Shigella* system revealed a direct interaction between $Spa40^{C}$ ($YscU^{C}$) and a 40 amino acid residues stretch within the proposed T3S4 (Type III secretion substrate specificity switch) region of Spa 32 (YscP) (Botteaux *et al*., 2008). Structural analysis would give further insights in the switching event.

During completion of this thesis the crystal structures of the cytoplasmic domain of FlhA and InvA (YscV homologues) have been determined (Saijo-Hamano *et al.*, 2010; Worrall *et al.*, 2010). Both structures are composed of a linker region and four unique sub domains. The sub domains of FlhA are proposed to shuttle between an open and closed conformation. The first sub domain resembles an ATPase binding domain of the circadian clock system of cyanobacteria suggesting an anchor point for the ATPase (Saijo-Hamano *et al.*, 2010). The first and second sub domain of InvA also exhibit structural similarity to the peripheral stalk of the A7V-type ATPase (Worrall *et al.*, 2010). Therefore structural investigations of the complex between the ATPase YscN and $YscV^{C}$ would provide deeper insights in the secretion process.

7 Bibliography

Abdallah, A. M., Gey van Pittius, N. C., Champion, P. A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C. M., Appelmelk, B. J. & Bitter, W. (2007). Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* **5**, 883-91.

Achtman, M., Zurth, K., Morelli, G., Torrea, G., Guiyoule, A. & Carniel, E. (1999). *Yersinia pestis*, the cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc Natl Acad Sci U S A* **96**, 14043-8.

Agrain, C., Sorg, I., Paroz, C. & Cornelis, G. R. (2005). Secretion of YscP from *Yersinia enterocolitica* is essential to control the length of the injectisome needle but not to change the type III secretion substrate specificity. *Mol Microbiol* **57**, 1415-27.

Akeda, Y. & Galan, J. E. (2005). Chaperone release and unfolding of substrates in type III secretion. *Nature* **437**, 911-5.

Allaoui, A., Woestyn, S., Sluiters, C. & Cornelis, G. R. (1994). YscU, a *Yersinia enterocolitica* inner membrane protein involved in Yop secretion. *J Bacteriol* **176**, 4534-42.

Ausubel, F. M., Brent, R., Kingston, R.E., Moor, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (2007). *Current protocols in molecular biology.*, John Wiley & Sons Inc., New York.

Bergfors, T. (2003). Seeds to crystals. *J Struct Biol* **142**, 66-76.

Bergsbaken, T. & Cookson, B. T. (2009). Innate immune response during *Yersinia* infection: critical modulation of cell death mechanisms through phagocyte activation. *J Leukoc Biol* **86**, 1153-8.

Blaylock, B., Riordan, K. E., Missiakas, D. M. & Schneewind, O. (2006). Characterization of the *Yersinia enterocolitica* type III secretion ATPase YscN and its regulator, YscL. *J Bacteriol* **188**, 3525-34.

Blocker, A., Gounon, P., Larquet, E., Niebuhr, K., Cabiaux, V., Parsot, C. & Sansonetti, P. (1999). The tripartite type III secreton of *Shigella flexneri* inserts IpaB and IpaC into host membranes. *J Cell Biol* **147**, 683-93.

Blocker, A., Jouihri, N., Larquet, E., Gounon, P., Ebel, F., Parsot, C., Sansonetti, P. & Allaoui, A. (2001). Structure and composition of the *Shigella flexneri* "needle complex", a part of its type III secreton. *Mol Microbiol* **39**, 652-63.

Botteaux, A., Sani, M., Kayath, C. A., Boekema, E. J. & Allaoui, A. (2008). Spa32 interaction with the inner-membrane Spa40 component of the type III secretion system of *Shigella flexneri* is required for the control of the needle length by a molecular tape measure mechanism. *Mol Microbiol* **70**, 1515-28.

Brunger, A. T. (1993). Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Crystallogr D Biol Crystallogr* **49**, 24-36.

Burghout, P., Beckers, F., de Wit, E., van Boxtel, R., Cornelis, G. R., Tommassen, J. & Koster, M. (2004a). Role of the pilot protein YscW in the biogenesis of the YscC secretin in *Yersinia enterocolitica*. *J Bacteriol* **186**, 5366-75.

Burghout, P., van Boxtel, R., Van Gelder, P., Ringler, P., Muller, S. A., Tommassen, J. & Koster, M. (2004b). Structure and electrophysiological properties of the

YscC secretin from the type III secretion system of *Yersinia enterocolitica*. *J Bacteriol* **186**, 4645-54.

Bürgi, H. B., Dunitz, J.D., Lehn, J.M. & Wipff, G. (1974). Stereochemistry of reaction paths at carbonyl centres. *Tetrahedron* **30**, 1563-1572.

Bürgi, H. B., Dunitz, J.D., Shefter, E. (1973). Geometrical reaction coordinates. II. Nucleophilic addition to a carbonyl group. *J Am Chem Soc* **95**, 5065-5067.

Chandran, V., Fronzes, R., Duquerroy, S., Cronin, N., Navaza, J. & Waksman, G. (2009). Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**, 1011-5.

Cianciotto, N. P. (2005). Type II secretion: a protein secretion system for all seasons. *Trends Microbiol* **13**, 581-8.

Coligan, J. E., Dunn, B.M., Ploegh, H.L., Speicher, D.W., and Wingfield, P.T. (2002). *Current protocols in protein science.*, John Wiley & Sons Inc., New York.

Cordes, F. S., Daniell, S., Kenjale, R., Saurya, S., Picking, W. L., Picking, W. D., Booy, F., Lea, S. M. & Blocker, A. (2005). Helical packing of needles from functionally altered *Shigella* type III secretion systems. *J Mol Biol* **354**, 206-11.

Cornelis, G. R. (2006). The type III secretion injectisome. *Nat Rev Microbiol* **4**, 811-25.

Cornelis, G. R., Boland, A., Boyd, A. P., Geuijen, C., Iriarte, M., Neyt, C., Sory, M. P. & Stainier, I. (1998). The virulence plasmid of *Yersinia*, an antihost genome. *Microbiol Mol Biol Rev* **62**, 1315-52.

Crepin, V. F., Shaw, R., Abe, C. M., Knutton, S. & Frankel, G. (2005). Polarity of enteropathogenic *Escherichia coli* EspA filament assembly and protein secretion. *J Bacteriol* **187**, 2881-9.

Dale, G. E., Oefner, C. & D'Arcy, A. (2003). The protein as a variable in protein crystallization. *J Struct Biol* **142**, 88-97.

Daniell, S. J., Takahashi, N., Wilson, R., Friedberg, D., Rosenshine, I., Booy, F. P., Shaw, R. K., Knutton, S., Frankel, G. & Aizawa, S. (2001). The filamentous type III secretion translocon of enteropathogenic *Escherichia coli*. *Cell Microbiol* **3**, 865-71.

Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-83.

Deane, J. E., Abrusci, P., Johnson, S. & Lea, S. M. (2010). Timing is everything: the regulation of type III secretion. *Cell Mol Life Sci* **67**, 1065-75.

Deane, J. E., Graham, S. C., Mitchell, E. P., Flot, D., Johnson, S. & Lea, S. M. (2008). Crystal structure of Spa40, the specificity switch for the *Shigella flexneri* type III secretion system. *Mol Microbiol* **69**, 267-76.

Deane, J. E., Roversi, P., Cordes, F. S., Johnson, S., Kenjale, R., Daniell, S., Booy, F., Picking, W. D., Picking, W. L., Blocker, A. J. & Lea, S. M. (2006). Molecular model of a type III secretion system needle: Implications for host-cell sensing. *Proc Natl Acad Sci U S A* **103**, 12529-33.

Delepelaire, P. (2004). Type I secretion in gram-negative bacteria. *Biochim Biophys Acta* **1694**, 149-61.

Dersch, P. (2003). Molecular and cellular mechanisms of bacterial entry into host cells. *Contrib Microbiol* **10**, 183-209.

Ding, Y., Xu, M. Q., Ghosh, I., Chen, X., Ferrandon, S., Lesage, G. & Rao, Z. (2003). Crystal structure of a mini-intein reveals a conserved catalytic module involved in side chain cyclization of asparagine during protein splicing. *J Biol Chem* **278**, 39133-42.

Dooley, C. T. & Houghten, R. A. (1993). The use of positional scanning synthetic peptide combinatorial libraries for the rapid determination of opioid receptor ligands. *Life Sci* **52**, 1509-17.

Edman, P. & Begg, G. (1967). A protein sequenator. *Eur J Biochem* **1**, 80-91.

Edqvist, P. J., Olsson, J., Lavander, M., Sundberg, L., Forsberg, A., Wolf-Watz, H. & Lloyd, S. A. (2003). YscP and YscU regulate substrate specificity of the *Yersinia* type III secretion system. *J Bacteriol* **185**, 2259-66.

Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.

Engh, R. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. . *Acta Crystallogr. Sect. A* **47**, 392-400.

Evans, G. & Pettifer, R. F. (2001). CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *Journal of Applied Crystallography* **34**, 82-86.

Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**, 72-82.

Ferris, H. U., Furukawa, Y., Minamino, T., Kroetz, M. B., Kihara, M., Namba, K. & Macnab, R. M. (2005). FlhB regulates ordered export of flagellar components via autocleavage mechanism. *J Biol Chem* **280**, 41236-42.

Frank, R. (2002). The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *J Immunol Methods* **267**, 13-26.

Fraser, G. M., Hirano, T., Ferris, H. U., Devgan, L. L., Kihara, M. & Macnab, R. M. (2003). Substrate specificity of type III flagellar protein export in Salmonella is controlled by subdomain interactions in FlhB. *Mol Microbiol* **48**, 1043-57.

French, G. S. & Wilson, K. S. (1978). On the treatment of negative intensity observations. *Acta Crystallogr A* **34**, 517-525.

Fronzes, R., Schafer, E., Wang, L., Saibil, H. R., Orlova, E. V. & Waksman, G. (2009). Structure of a type IV secretion system core complex. *Science* **323**, 266-8.

Gal, M., Schanda, P., Brutscher, B. & Frydman, L. (2007). UltraSOFAST HMQC NMR and the repetitive acquisition of 2D protein spectra at Hz rates. *J Am Chem Soc* **129**, 1372-7.

Galan, J. E. & Wolf-Watz, H. (2006). Protein delivery into eukaryotic cells by type III secretion machines. *Nature* **444**, 567-73.

Garman, E. F. & Owen, R. L. (2006). Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **62**, 32-47.

Giddings, J. C. (1993). Field-flow fractionation: analysis of macromolecular, colloidal, and particulate materials. *Science* **260**, 1456-65.

Gonzalez-Pedrajo, B., Fraser, G. M., Minamino, T. & Macnab, R. M. (2002). Molecular dissection of *Salmonella* FliH, a regulator of the ATPase FliI and the type III flagellar protein export pathway. *Mol Microbiol* **45**, 967-82.

Gophna, U., Ron, E. Z. & Graur, D. (2003). Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. *Gene* **312**, 151-63.

Guerrero, S. A., Hecht, H. J., Hofmann, B., Biebl, H. & Singh, M. (2001). Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. *Appl Microbiol Biotechnol* **56**, 718-23.

Gunasekaran, K., Ramakrishnan, C. & Balaram, P. (1997). Beta-hairpins in proteins revisited: lessons for *de novo* design. *Protein Eng* **10**, 1131-41.

Guntert, P., Dotsch, V., Wider, G. & Wuthrich, K. (1992). Processing of Multidimensional Nmr Data with the New Software Prosa. *Journal of Biomolecular Nmr* **2**, 619-629.

Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. & Ala'Aldeen, D. (2004). Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* **68**, 692-744.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *Embo J* **9**, 1665-72.

Hendrickson, W. A., Smith, J. L. & Sheriff, S. (1985). Direct phase determination based on anomalous scattering. *Methods Enzymol* **115**, 41-55.

Hodgkinson, J. L., Horsley, A., Stabat, D., Simon, M., Johnson, S., da Fonseca, P. C., Morris, E. P., Wall, J. S., Lea, S. M. & Blocker, A. J. (2009). Three-dimensional reconstruction of the *Shigella* T3SS transmembrane regions reveals 12-fold symmetry and novel features throughout. *Nat Struct Mol Biol* **16**, 477-85.

Hoiczyk, E. & Blobel, G. (2001). Polymerization of a single protein of the pathogen *Yersinia enterocolitica* into needles punctures eukaryotic cells. *Proc Natl Acad Sci U S A* **98**, 4669-74.

Isberg, R. R. & Barnes, P. (2001). Subversion of integrins by enteropathogenic *Yersinia*. *J Cell Sci* **114**, 21-28.

Journet, L., Agrain, C., Broz, P. & Cornelis, G. R. (2003). The needle length of bacterial injectisomes is determined by a molecular ruler. *Science* **302**, 1757-60.

Kall, L., Krogh, A. & Sonnhammer, E. L. (2004). A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* **338**, 1027-36.

Kantardjieff, K. A. & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci* **12**, 1865-71.

Keller, R. (2004). Optimizing the process of NMR spectrum analysis and computer aided resonance assignment, ETH Zürich.

Knutton, S., Rosenshine, I., Pallen, M. J., Nisan, I., Neves, B. C., Bain, C., Wolff, C., Dougan, G. & Frankel, G. (1998). A novel EspA-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. *Embo J* **17**, 2166-76.

Koster, M., Bitter, W., de Cock, H., Allaoui, A., Cornelis, G. R. & Tommassen, J. (1997). The outer membrane component, YscC, of the Yop secretion machinery of *Yersinia enterocolitica* forms a ring-shaped multimeric complex. *Mol Microbiol* **26**, 789-97.

Krissinel, E. & Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-97.

Kubori, T., Matsushima, Y., Nakamura, D., Uralil, J., Lara-Tejero, M., Sukhan, A., Galan, J. E. & Aizawa, S. I. (1998). Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. *Science* **280**, 602-5.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.

Lavander, M., Sundberg, L., Edqvist, P. J., Lloyd, S. A., Wolf-Watz, H. & Forsberg, A. (2002). Proteolytic cleavage of the FlhB homologue YscU of *Yersinia pseudotuberculosis* is essential for bacterial survival but not for type III secretion. *J Bacteriol* **184**, 4500-9.

Lavander, M., Sundberg, L., Edqvist, P. J., Lloyd, S. A., Wolf-Watz, H. & Forsberg, A. (2003). Characterisation of the type III secretion protein YscU in *Yersinia pseudotuberculosis*. YscU cleavage--dispensable for TTSS but essential for survival. *Adv Exp Med Biol* **529**, 109-12.

Leslie, A. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26**.

Lottspeich, F. & Engels, J. W. (2006). *Bioanalytik*, 2, Elsevier, Spektrum Akademischer Verlag Munich.

Lountos, G. T., Austin, B. P., Nallamsetty, S. & Waugh, D. S. (2009). Atomic resolution structure of the cytoplasmic domain of *Yersinia pestis* YscU, a regulatory switch involved in type III secretion. *Protein Sci* **18**, 467-74.

Lovell, S. C., Davis, I. W., Arendall, W. B. I., Bakker, P. I. D., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). Structure validation by C-alpha geometry: phi, psi and C-beta deviation. *Proteins* **50, 437-450.**, 437-450.

Luft, J. R. & DeTitta, G. T. (1997). Kinetic aspects of macromolecular crystallization. *Methods Enzymol* **276**, 110-31.

Macnab, R. M. (2003). How bacteria assemble flagella. *Annu Rev Microbiol* **57**, 77-100.

Macnab, R. M. (2004). Type III flagellar protein export and flagellar assembly. *Biochim Biophys Acta* **1694**, 207-17.

Marlovits, T. C., Kubori, T., Lara-Tejero, M., Thomas, D., Unger, V. M. & Galan, J. E. (2006). Assembly of the inner rod determines needle length in the type III secretion injectisome. *Nature* **441**, 637-40.

Marlovits, T. C., Kubori, T., Sukhan, A., Thomas, D. R., Galan, J. E. & Unger, V. M. (2004). Structural insights into the assembly of the type III secretion needle complex. *Science* **306**, 1040-2.

Marlovits, T. C. & Stebbins, C. E. (2010). Type III secretion systems shape up as they ship out. *Curr Opin Microbiol* **13**, 47-52.

Matthews, B. W. (1968). Solvent content of protein crystals. *J Mol Biol* **33**, 491-7.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* **61**, 458-64.

McDonald, C., Vacratsis, P. O., Bliska, J. B. & Dixon, J. E. (2003). The *yersinia* virulence factor YopM forms a novel protein complex with two cellular kinases. *J Biol Chem* **278**, 18514-23.

McMurry, J. L., Van Arnam, J. S., Kihara, M. & Macnab, R. M. (2004). Analysis of the cytoplasmic domains of *Salmonella* FlhA and interactions with components of the flagellar export machinery. *J Bacteriol* **186**, 7586-92.

Minamino, T., Imada, K. & Namba, K. (2008). Mechanisms of type III protein export for bacterial flagellar assembly. *Mol Biosyst* **4**, 1105-15.

Minamino, T. & Macnab, R. M. (2000a). Domain structure of *Salmonella* FlhB, a flagellar export component responsible for substrate specificity switching. *J Bacteriol* **182**, 4906-14.

Minamino, T. & MacNab, R. M. (2000b). Interactions among components of the *Salmonella* flagellar export apparatus and its substrates. *Mol Microbiol* **35**, 1052-64.

Minamino, T., Shimada, M., Okabe, M., Saijo-Hamano, Y., Imada, K., Kihara, M. & Namba, K. (2010). Role of the C-terminal cytoplasmic domain of FlhA in bacterial flagellar type III protein export. *J Bacteriol*, in press.

Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.* **374**, 229-244.

Mueller, C. A., Broz, P., Muller, S. A., Ringler, P., Erne-Brand, F., Sorg, I., Kuhn, M., Engel, A. & Cornelis, G. R. (2005). The V-antigen of *Yersinia* forms a distinct structure at the tip of injectisome needles. *Science* **310**, 674-6.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55.

Navarro, L., Koller, A., Nordfelth, R., Wolf-Watz, H., Taylor, S. & Dixon, J. E. (2007). Identification of a molecular target for the *Yersinia* protein kinase A. *Mol Cell* **26**, 465-77.

Niesen, F. H., Berglund, H. & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2**, 2212-21.

Okon, M., Moraes, T. F., Lario, P. I., Creagh, A. L., Haynes, C. A., Strynadka, N. C. & McIntosh, L. P. (2008). Structural characterization of the type-III pilot-secretin complex from *Shigella flexneri*. *Structure* **16**, 1544-54.

Otwinowski, Z. M., W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.

Pallen, M. J., Beatson, S. A. & Bailey, C. M. (2005). Bioinformatics, genomics and evolution of non-flagellar type-III secretion systems: a Darwinian perspective. *FEMS Microbiol Rev* **29**, 201-29.

Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr* **61**, 449-57.

Pape, T. & Schneider, T. R. (2004). HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Cryst.* **37**, 843-844.

Perry, R. D. & Fetherston, J. D. (1997). *Yersinia pestis*--etiologic agent of plague. *Clin Microbiol Rev* **10**, 35-66.

Podzimek, S., Lebeda, P., Johann, C. & Jocks, T. (2008). Asymmetrische Fluss-Feldfluss-Fraktionierung - eine leistungsfähige Methode zur Polymer-charakterisierung. doi: http://www.chemie.de/articles/d/108885.

Pozidis, C., Chalkiadaki, A., Gomez-Serrano, A., Stahlberg, H., Brown, I., Tampakaki, A. P., Lustig, A., Sianidis, G., Politou, A. S., Engel, A., Panopoulos, N. J., Mansfield, J., Pugsley, A. P., Karamanou, S. & Economou, A. (2003). Type III protein translocase: HrcN is a peripheral ATPase that is activated by oligomerization. *J Biol Chem* **278**, 25816-24.

Prentice, M. B. & Rahalison, L. (2007). Plague. *Lancet* **369**, 1196-207.

Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Adv Protein Chem* **23**, 283-438.

Rhodes, G. (2006). *Crystallography Made Crystal Clear*. 3 edit, Elsevier Inc.

Riedel, E. & Janiak, C. (2002). *Anorganische Chemie*, 5, de Gruyter, Berlin; New York.

Riordan, K. E. & Schneewind, O. (2008). YscU cleavage and the assembly of *Yersinia* type III secretion machine complexes. *Mol Microbiol* **68**, 1485-501.

Rollins, S. E., Rollins, S. M. & Ryan, E. T. (2003). *Yersinia pestis* and the plague. *Am J Clin Pathol* **119 Suppl**, S78-85.

Ruckdeschel, K., Mannel, O., Richter, K., Jacobi, C. A., Trulzsch, K., Rouot, B. & Heesemann, J. (2001). *Yersinia* outer protein P of *Yersinia enterocolitica* simultaneously blocks the nuclear factor-kappa B pathway and exploits lipopolysaccharide signaling to trigger apoptosis in macrophages. *J Immunol* **166**, 1823-31.

Rupp, B. (2006). Real-space solution to the problem of full disclosure. *Nature* **444**, 817.

Saijo-Hamano, Y., Imada, K., Minamino, T., Kihara, M., Macnab, R. M. & Namba, K. (2005). Crystallization and preliminary X-ray analysis of the C-terminal

cytoplasmic domain of FlhA, a membrane-protein subunit of the bacterial flagellar type III protein-export apparatus. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **61**, 599-602.

Saijo-Hamano, Y., Imada, K., Minamino, T., Kihara, M., Shimada, M., Kitao, A. & Namba, K. (2010). Structure of the cytoplasmic domain of FlhA and implication for flagellar type III protein export. *Mol Microbiol*.

Saijo-Hamano, Y., Minamino, T., Macnab, R. M. & Namba, K. (2004). Structural and functional analysis of the C-terminal cytoplasmic domain of FlhA, an integral membrane component of the type III flagellar protein export apparatus in *Salmonella*. *J Mol Biol* **343**, 457-66.

Sambrook, J. a. R., D.W. (2000). *Molecular Cloning - a laboratory manual.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Schneider, T. R. & Sheldrick, G. M. (2002). Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772-1779.

Shao, F. (2008). Biochemical functions of *Yersinia* type III effectors. *Curr Opin Microbiol* **11**, 21-9.

Shao, F., Vacratsis, P. O., Bao, Z., Bowers, K. E., Fierke, C. A. & Dixon, J. E. (2003). Biochemical characterization of the *Yersinia* YopT protease: cleavage site and recognition elements in Rho GTPases. *Proc Natl Acad Sci U S A* **100**, 904-9.

Shrivastava, S. & Mande, S. S. (2008). Identification and functional characterization of gene components of Type VI Secretion system in bacterial genomes. *PLoS One* **3**, e2955.

Sorg, I., Wagner, S., Amstutz, M., Muller, S. A., Broz, P., Lussi, Y., Engel, A. & Cornelis, G. R. (2007). YscU recognizes translocators as export substrates of the *Yersinia* injectisome. *EMBO J* **26**, 3015-24.

Spreter, T., Yip, C. K., Sanowar, S., Andre, I., Kimbrough, T. G., Vuckovic, M., Pfuetzner, R. A., Deng, W., Yu, A. C., Finlay, B. B., Baker, D., Miller, S. I. & Strynadka, N. C. (2009). A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nat Struct Mol Biol* **16**, 468-76.

Tamano, K., Aizawa, S., Katayama, E., Nonaka, T., Imajoh-Ohmi, S., Kuwae, A., Nagai, S. & Sasakawa, C. (2000). Supramolecular structure of the *Shigella* type III secretion machinery: the needle part is changeable in length and essential for delivery of effectors. *EMBO J* **19**, 3876-87.

Tong, L. & Rossmann, M. G. (1997). Rotation function calculations with GLRF program. *Methods Enzymol* **276**, 594-611.

Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4.

Troisfontaines, P. & Cornelis, G. R. (2005). Type III secretion: more systems than you think. *Physiology (Bethesda)* **20**, 326-39.

Trosky, J. E., Liverman, A. D. & Orth, K. (2008). *Yersinia* outer proteins: Yops. *Cell Microbiol* **10**, 557-65.

Trulzsch, K., Oellerich, M. F. & Heesemann, J. (2007). Invasion and dissemination of *Yersinia enterocolitica* in the mouse infection model. *Adv Exp Med Biol* **603**, 279-85.

Tseng, T. T., Tyler, B. M. & Setubal, J. C. (2009). Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol*, S2.

Van Arnam, J. S., McMurry, J. L., Kihara, M. & Macnab, R. M. (2004). Analysis of an engineered *Salmonella* flagellar fusion protein, FliR-FlhB. *J Bacteriol* **186**, 2495-8.

Van Gijsegem, F., Vasse, J., Camus, J. C., Marenda, M. & Boucher, C. (2000). *Ralstonia solanacearum* produces hrp-dependent pili that are required for PopA secretion but not for attachment of bacteria to plant cells. *Mol Microbiol* **36**, 249-60.

Viboud, G. I. & Bliska, J. B. (2005). *Yersinia* outer proteins: role in modulation of host cell signaling responses and pathogenesis. *Annu Rev Microbiol* **59**, 69-89.

Von Pawel-Rammingen, U., Telepnev, M. V., Schmidt, G., Aktories, K., Wolf-Watz, H. & Rosqvist, R. (2000). GAP activity of the *Yersinia* YopE cytotoxin specifically targets the Rho pathway: a mechanism for disruption of actin microfilament structure. *Mol Microbiol* **36**, 737-48.

Wagner, S., Sorg, I., Degiacomi, M., Journet, L., Dal Peraro, M. & Cornelis, G. R. (2009). The helical content of the YscP molecular ruler determines the length of the *Yersinia* injectisome. *Mol Microbiol* **71**, 692-701.

Wiesand, U., Sorg, I., Amstutz, M., Wagner, S., van den Heuvel, J., Luhrs, T., Cornelis, G. R. & Heinz, D. W. (2009). Structure of the type III secretion recognition protein YscU from *Yersinia enterocolitica*. *J Mol Biol* **385**, 854-66.

Wilharm, G., Lehmann, V., Krauss, K., Lehnert, B., Richter, S., Ruckdeschel, K., Heesemann, J. & Trulzsch, K. (2004). *Yersinia enterocolitica* type III secretion depends on the proton motive force but not on the flagellar motor components MotA and MotB. *Infect Immun* **72**, 4004-9.

Wilson, A. J. C. (1949). The probability distributions of X-ray intensities. *Acta Crystallogr.* **2**, 318-321.

Woestyn, S., Allaoui, A., Wattiau, P. & Cornelis, G. R. (1994). YscN, the putative energizer of the *Yersinia* Yop secretion machinery. *J Bacteriol* **176**, 1561-9.

Wood, S. E., Jin, J. & Lloyd, S. A. (2008). YscP and YscU switch the substrate specificity of the *Yersinia* type III secretion system by regulating export of the inner rod protein YscI. *J Bacteriol* **190**, 4252-62.

Wooldridge, K. (2009). *Bacterial Secreted Proteins: Secretory Mechanisms and Role in Pathogenesis.*, Caister Academic Press.

Worrall, L. J., Vuckovic, M. & Strynadka, N. C. (2010). Crystal structure of the C-terminal domain of the *Salmonella* type III secretion system export apparatus protein InvA. *Protein Sci*, in press.

Wren, B. W. (2003). The *yersiniae*--a model genus to study the rapid evolution of bacterial pathogens. *Nat Rev Microbiol* **1**, 55-64.

Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids.*, John Wiley & Sons, New York.

Yip, C. K., Kimbrough, T. G., Felise, H. B., Vuckovic, M., Thomas, N. A., Pfuetzner, R. A., Frey, E. A., Finlay, B. B., Miller, S. I. & Strynadka, N. C. (2005). Structural characterization of the molecular platform for type III secretion system assembly. *Nature* **435**, 702-7.

Young, G. M., Schmiel, D. H. & Miller, V. L. (1999). A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proc Natl Acad Sci U S A* **96**, 6456-61.

Zarivach, R., Deng, W., Vuckovic, M., Felise, H. B., Nguyen, H. V., Miller, S. I., Finlay, B. B. & Strynadka, N. C. (2008). Structural analysis of the essential self-cleaving type III secretion proteins EscU and SpaS. *Nature* **453**, 124-7.

Zarivach, R., Vuckovic, M., Deng, W., Finlay, B. B. & Strynadka, N. C. (2007). Structural analysis of a prototypical ATPase from the type III secretion system. *Nat Struct Mol Biol* **14**, 131-7.

Zhu, K., Gonzalez-Pedrajo, B. & Macnab, R. M. (2002). Interactions among membrane and soluble components of the flagellar export apparatus of *Salmonella*. *Biochemistry* **41**, 9516-24.

8 Overview: *Yersinia* injectisome

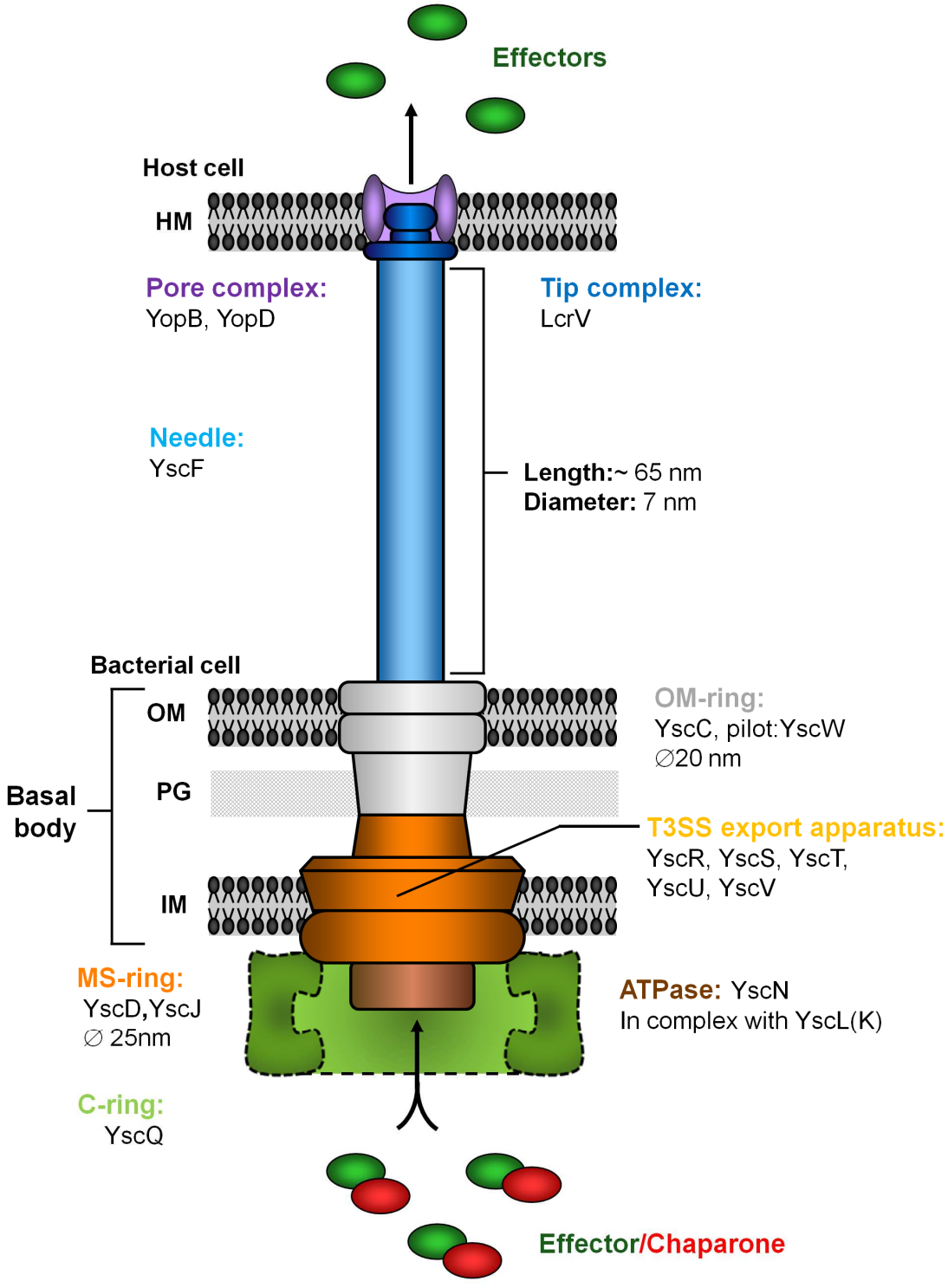


Figure 55: Schematic representation of the *Yersinia* injectisome. HM, host cell membrane; IM, inner membrane; OM, outer membrane; PG, peptidoglycan. The C-ring is represented in dashed lines as information on this component is still scare. Colour code of different structural components corresponds to Ysc/Yop proteins in Table 12.

Table 12: Overview about *Yersinia* Injectisome proteins and their homolog's. (Please note: This is not a complete list of all proteins involved in the T3SS!)

Ysc/Yop	kDa	Function	Contact	T3S Homologues	Flagellar	Structure
Structural Components of the Injectisome						
YopB	41.9	Translocator, pore	YopD, LcrV	SipB[1], IpaB[2], EspD[3]	None	No
YopD	33.2	Translocator, pore	YopB, LcrV	SipC[1], IpaC[2], EspB[3]	None	No
LcrV	37.5	Needle tip structure	YopB, YopD	SipD[1], IpaD[2], EspA[3]	None	LcrV
YscF	9.4	Needle subunit	LcrV	PrgI[1], MxiH[2], EscF[3]	-	MxiH, YscF
YscC	67.1	Secretin, OM ring	YscW, YscD	InvG[1], MxiD[2], EscC[3]	none	EscC N-ter.
YscD	46.7	Upper part of MS ring	YscC, YscJ	PrgH[1], MxiG[2], EscD[3]	-	PrgH C-ter.
YscJ	27.0	Lowerof the MS ring	YscD	PrgK[1], MxiJ[2], EscJ[3]	FliF	EscJ
YscR	24.4	Export apparatus	YscS, YscU	InvL[1], Spa24[2], EscR[3]	FliP	No
YscS	9.6	Export apparatus	YscR	SpaQ[1], Spa9[2], EscS[3]	FliQ	No
YscT	28.4	Export apparatus	-	InvN[1], Spa29[2], EscT[3]	FliR	No
YscU	40.3	Export apparatus	YscR,YscV, N-ter YscP	SpaS[1], Spa40[2], EscU[3]	FlhB	C-ter: Spa40 EscU, YscU
YscV	77.8	Export apparatus	YscU	InvA[1], MxiA[2], EscV[3]	FlhA	No
YscK	23.9	Docking of the ATPase to the C-ring	YscQ	MxiK[2]	-	No
YscL	24.9	ATPase Regulator	YscN, YscQ	MxiN[2]	FliH	No
YscN	47.8	ATPase	YscL, YscK	InvC[1], Spa47[2], EscN[3]	FliI	FliI, EscN
YscQ	34.4	Subunit of the C-ring	YscK, YscL	SpaO/InvK[1] ,Spa33[2]	FliN + FliM	FliN/M, HrpQ[4]
Important Proteins for Assembly						
YscP	57.4	Molecular ruler and substrate specificity switch	N-term with YscU C-ter	InvJ[1], Spa32[2]	FliK	No
YscW	14.6	Pilotin for YscC	YscC	InvH[1], MxiM[2]	-	MxiM PilP
YscH	18	Early export substrate. Not needed for needle assembly	-	PscH[5]	-	No
YscI	12.7	Early export substrate. Needed for needle assembly.	-	-	-	No
YscO	19	Substrate export. Needed for needle assembly/T3S.	-	CT670[6]	-	CT670

Proteins listed are: 1, *Salmonella thypimurium*; 2, *Shigella spp.*; 3, *Enteropathogenic Escherichia coli*; 4, *Pseudomonas syringae*; 5, *Pseudomonas aeruginosa*; 6, *Chlamydia trachomatis*

Danksagung

Prof. Dr. Dirk Heinz danke ich herzlich für die Aufnahme in seine Arbeitsgruppe und der Möglichkeit meine Promotionsarbeit in dem spannenden Feld des T3SS unter hervorragenden Bedingungen durchführen zu können. Sein Interesse am Fortgang meiner Arbeit sowie seine Diskussionsbereitschaft halfen maßgeblich zum Erlangen dieser Ergebnisse bei. Desweiteren möchte ich mich dafür bedanken, dass ich an zahlreichen interessanten Tagungen und Konferenzen teilnehmen und meine Ergebnissen präsentieren durfte, so dass ich meinen wissenschaftlichen Horizont erheblich erweitern konnte.

Prof. Dr. Michael Steinert danke ich sehr herzlich für die bereitwillige Übernahme des Zweitgutachtens.

Bei Prof. Dr. Dieter Jahn bedanke ich mich herzlich für die Übernahme des Prüfungsvorsitzes.

Bei Prof. Dr. Guy Cornelis und Dr. Isabel Sorg sowie dem ganzen “*Yersinia* Lab” aus Basel möchte ich mich für die hervorragende Kooperation und ständige Diskussionsbereitschaft bedanken.

Dr. Victor Wray und Dr. Wolf-Dieter Schubert möchte ich ganz herzlich für ihre Hilfsbereitschaft beim “*paper submitten*“ danken. Außerdem hatten sie stets ein offenes Ohr für Fragen aller Art.

Bei Dr. Maike Bublitz, Dr. Maike Rochon, Agnes Zimmer und Lilia Polle möchte ich besonders für die Starthilfe am Anfang meiner Doktorarbeit bedanken.

Bei Kevin Walkling möchte ich mich für seine wertvolle Hilfe und fortwährende Unterstützung bei Präparationen und Kristallisationen diverser Proteine bedanken.

Meinen Diplomanden Hauke Pagel und Stefan Münnich möchte ich für ihr großes Engagement im Labor und den tollen Ergebnissen danken.

Dr. Thorsten Lührs danke ich für die Unterstützung und Umsetzungen von NMR und AF4 Experimenten.

Allen Mitgliedern der SB gilt mein Dank für ein freundliches Arbeitsklima und den vielen kleinen und großen Unterstützungen, die hier nicht alle erwähnt werden können.

Zum Schluss möchte ich mich ganz besonders bei meiner Freundin Malina und meiner Familie bedanken, die mich trotz der weiten Entfernung immer wieder aufgemuntert haben. Vielen lieben Dank dafür!

Lebenslauf

Persönliche Angaben

Ulrich Andreas Philipp Wiesand
geb. am 21.07.1981 in Füssen
ledig
deutsch

Schulischer Werdegang

1987-1991	Grundschule Rieden a.F.
1991-1992	Anton-Sturm-Volksschule Füssen
1992-2001	Math.-naturwissenschatliches Gymnasium Hohenschwangau Abschluss: Abitur

Akademischer Werdegang

2001-2007	Studium der Chemie an der Albert-Ludwigs-Universität Freiburg Diplomarbeit in der Arbeitsgruppe von Prof. Georg E. Schulz
2007-2010	Promotion am Helmholtz Zentrum für Infektionsforschung im Bereich Strukturbiologie in der Arbeitsgruppe von Prof. Dirk W. Heinz.